مجمع الملك سلمان العالمي للغة العربية
King Salman Global Academy for Arabic Language

السنة الأولى - العدد الأول
جمادى الأول ١٤٤٥هـ - ديسمبر ٢٠٢٣م

اختبارات اللغة

Journal of Language Testing

مقارنة نتائج الاختبارات المحوسبة بنتائج نظيراتها الورقية في الاختبارات اللغوية

استعراض لاختبار المستوى الرابع لتخصص اللغة العربية في الجامعات الصينية: واقعه والمأمول منه.

دراسةٌ مقارنة بين توجيهاتِ المجلس الأمريكي لتعليم اللغات الأجنبية والإطارِ الأوربيّ المرجعيّ المشترك للُغات.

الاختبارات الموضوعية الأكثر استخدامًا والمستويات والنواتج التي تستهدفها.

تصور مقترح لبناء اختبار معياري لقياس كفايات معلمي اللغة العربية لغة ثانية.

How Different Are Test-Taking Strategies with Computer-Based and Paper-Based Reading? The Case of Saudi EFL Learners.

مجلة علمية مُحكمة تصدر عن مجمع الملك سلمان العالمي للغة العربية
www.ksaa.gov.sa

مجمع الملك سلمان العالمي للغة العربية
King Salman Global Academy for Arabic Language

السنة الأولى - العدد الأول
جمادى الأول ١٤٤٥ هـ - ديسمبر ٢٠٢٣ م

اختبارات اللغة

Journal of Language Testing
www.ksaa.gov.sa
مجلة علمية مُحكمة تصدر عن مجمع الملك سلمان العالمي للغة العربية

مجمع الملك سلمان العالمي للغة العربية
King Salman Global Academy for Arabic Language

المشرف العام على المجلة

أ.د عبدالله بن صالح الوشمي

إدارة التحرير

أ.د خالد بن عبدالعزيز الدامغ
رئيس التحرير

د. خالد بن عبدالله الصقير
مدير التحرير

أ. سارة بنت محمد السريّع
أمينة المجلة

هيئة التحرير

أ.د حسن بن محمد الشمراني
أ.د سعيد بن عبدالله الدوسري
د. عبدالرحمن بن هادي الشمراني
د. عبدالله بن عبدالوهاب الفريدان
أ.د وليد بن أحمد العناتي

الهيئة الاستشارية

د. أحمد السقوفي (السعودية)
أ.د. خالد جمال المحمدي (العراق)
أ.د. خليل الحربي (السعودية)
أ.د. راشد المحرزي (عمان)
د. محمد صلاح الدين سالم (مصر)
د. محمد عبدالخالق (السودان)
أ.د. محمد كنتاوي (الجزائر)
أ.د هنادا طه (لبنان)

مجمع الملك سلمان العالمي للغة العربية
King Salman Global Academy for Arabic Language

لإسهاماتكم واستفساراتكم، نسعد بتواصلكم من خلال:

البريد الإلكتروني: jlt@ksaa.gov.sa

هاتف: 00966 112587269 00966 112581069

ص.ب ١٢٥٠٠ الرياض ١١٤٧٣

المملكة العربية السعودية

www.ksaa.gov.sa

صفة المجلة

مجلة اختبارات اللغة هي مجلة علمية محكمة تصدر عن مجمع الملك سلمان العالمي للغة العربية بدءا من عام (١٤٤٥هـ / ٢٠٢٣) وهي متخصصة في مجال التقييم اللغوي، وتشمل اختبارات اللغة، والتقييم اللغوي، والقياس، وما يتبعها من تفرعات.

قواعد النشر

- أن يكون البحث مكتوباً بلغة عربية سليمة، وتقبل الأبحاث باللغة الإنجليزية.
- أن يكون البحث في أحد مجالات الاختبارات والمقاييس اللغوية.
- أن يعتمد الباحث الأسس العلمية في جميع خطواته.
- أن يتبع قواعد التوثيق العلمي في المتن وفي المراجع (والطريقة المتعبة هي طريقة جمعية علم النفس الأمريكية APA).
- ألا يقل البحث أو الدراسة عن ٢٠٠٠ كلمة، ولا يزيد عن ١٠ آلاف.
- ألا يكون قد سبق نشر البحث في أي وعاء علمي.
- أن يعد الباحث ملخصا لبحثه في حدود ٢٥٠ كلمة باللغتين العربية والإنجليزية.

المجالات البحثية للمجلة

ينطلق اهتمام المجلة بالبحوث المتعلقة بمجالات الاختبارات اللغوية لأبناء العربية ومتعلميها، بما في ذلك الدراسات المتعلقة بمعايير اختبارات الكفاية اللغوية، ومصطلحات الاختبارات ومفاهيمها، وصياغة اختبارات الكفاية، وأدوات قياس الكفاية اللغوية وآلياتها.

اختبارات اللغة
Journal of Language Testing

بيانات المجلة:

مجلة اختبارات اللغة
مجلة علمية محكمة تصدر عن
مجمع الملك سلمان العالمي للغة العربية
الرياض – المملكة العربية السعودية
ردمد النسخة الورقية: ١٦٥٨-٩٤٩١
ردمد النسخة الإلكتروني: ١٦٥٨-٩٥٠٥

Journal information:

Journal of Language Testing
A peer-reviewed bi-annual journal issued by
King Salman Global Academy for Arabic Language
Riyadh- Saudi Arabia
ISBN paper version: 9491-1658
ISBN Electronic version: 9505-1658

محتويات العدد

مقارنة نتائج الاختبارات المحوسبة بنتائج نظيراتها الورقية في الاختبارات اللغوية
أ.د: خالد بن عبدالعزيز الدامغ
أ. سارة بنت عبدالله القحطاني
٧

استعراض لاختبار المستوى الرابع لتخصص اللغة العربية في الجامعات الصينية: واقعه والمأمول منه
د: يه ليانغ ينغ
٥١

دراسةٌ مقارنة بين توجيهاتِ المجلس الأمريكي لتعليم اللغات الأجنبية والإطارِ الأوربيّ المرجعيّ المشترك للُّغات
د.إسلام يسري الحدقي
٨٠

الاختبارات الموضوعية الأكثر استخدامًا والمستويات والنواتج التي تستهدفها
د. محمد عبد الخالق محمد فضل
١١٦

تصور مقترح لبناء اختبار معياري لقياس كفايات معلمي اللغة العربية لغة ثانية
د:عبيد الله بن عبدالله الفايدي
١٦٧

How Different Are Test-Taking Strategies with Computer-Based and Paper-Based Reading? The Case of Saudi EFL Learners
د.عبدالعزيز بن عبدالرحمن أبانمي
٢٠٧

كلمة المجمع

ضمن جهود المملكة العربية السعودية في رفع مكانة اللغة العربية، ودعم حضورها العالمي والعلمي؛ يأتي تأسيس مجمع الملك سلمان العالمي للغة العربية؛ لتعزيز مكانة اللغة العربية، ومواصلة المسير في دعمها والحفاظ عليها، ويحظى المجمع بإشراف وقيادة سمو وزير الثقافة رئيس مجلس أمناء المجمع الأمير بدر بن عبدالله بن فرحان آل سعود حفظه الله، وينطلق في تنفيذ برامجه التعليمية وفعالياته الثقافية داخل العالم العربي وخارجه، والعمل الوثيق مع المنظمات الدولية والمؤسسات اللغوية.

ولا يغفل المجمع ضمن نشاطه العالمي عن العمق اللغوي والمعرفي؛ فبه يشتد عود المعرفة، ويقوى نشاطها، ومن أبرز اهتمامات المجمع العمل في دعم الحراك العلمي اللغوي، ومن ذلك إنشاء المجلات العلمية النوعية؛ فقد أطلق المجمع ثلاث مجلات علمية جديدة تساند المجلات والمجالات السابقة في خدمة اللغة العربية، وهي مجلة اللغويات الحاسوبية والمعالجة الآلية للغة العربية، ومجلة علوم اللغة العربية، ومجلة اختبارات اللغة.

وتأتي أهمية مجلة اختبارات اللغة لتواكب مشـــروع المجمع العالمي الكبير في إطلاق اختباره اللغوي الدولي، ولندرة البحوث النوعية في هذا الســـياق، والحاجة المســـتمرة إلى تجويد البحث والمعـــرف في حقل اختبارات اللغة في العالـــم، وفي العالم العربي على وجه أخص؛ فالاختبارات اللغوية حجر زاوية في التعليم اللغوي، وانعكاســـاتها تتصل بجميع أجـــزاء العمليـــة التعليمية؛ فهي تؤثـــر في المناهج العلميـــة وتطويرهـــا، وتقويم التعلم والتعليم والمتعلم.

أثـــق بأن يكون هذا العدد محفزًا على حراك علمي وبحثي في ســـياق القياس والتقويم اللغـــوي يخدم اللغة العربيـــة وتعليمها وتعلمها، وأشـــيد بجهود هيئة التحرير، وأشـــكر للهيئة الاستشـــارية تعاونهم، ومما يســـعدني وزملائي فريق المجلة استقبال مشاركات علمية رفيعة القدر من العلماء والباحثين لإثراء هذا الحقل العلمي المتخصص بإنتاجهم العلميّ والبحثيّ.

الأمين العام
المشرف العام على المجلة
أ.د عبد الله بن صالح الوشمي

هذه الطبعة **إهـداء من المجمـع**
ولا يُسمح بنشرها ورقياً أو تداولها تجارياً

كلمة التحرير

تسـنَّمت المؤسسـات العلميـة في المملكـة العربيـة السـعودية نظيراتهـا في ميادين علـم اللغـة التطبيقـي عامـة، وفي اللغويـات التطبيقيـة العربيـة خاصـة، حيـث تتميـز مؤسسـاتها في هـذا المجـال برصانـة البرامـج العلميـة وعلـو الأداء المهـني ونوعيـة المشـاركات العالميـة في المجـالات اللغويـة التطبيقيـة المختلفـة. وامتـداداً لهـذا التميـز تنطلـق هـذه المجلـة العلميـة لتملأ فراغـاً، وتلبي حاجـة، وتبـني علمـا في مجـال القيـاس والتقييـم اللغـوي. ويسـعدنا في هـذا الصـدد أن نقـدم باكـورة إصـدارات هـذه المجلـة المتخصصـة في «الاختبـارات اللغويـة» راجـين أن تكـون وعـاء نشـر فاعـل للتجسـير المعـرفي بـين المهتمـين بالاختبـارات، وأن تكـون نبراسـاً ينهـل منـه المتخصصـون وطلاب العلـم، خاصـة وأنهـا الأولى في مجالهـا في الوطـن العـربي، ومـن المجـلات العلميـة القلائـل المتخصصـة في مجـال الاختبـارات علـى مسـتوى العالـم.

وإذ نقـدم العـدد الأول مـن مجلـة «اختبـارات اللغـة» فإننـا علـى قناعـة بـأن الأبحـاث الـتي تضمنهـا العـدد الأول تتسـم بوجهـي العملـة كليهمـا، وجـه المحتـوى العلـمي الثمـين، ووجـه البنـاء البحـثي الرصـين، وهمـا وجهـان يتكاملان للـرقي بالبحـث العلـمي. وقـد تضمـن عقـد البحـوث في هـذا العـدد السلسـلة التاليـة مـن الدراسـات:

البحـث الأول تنـاول قضيـة تتصـدر اهتمامـات المعنيـين بالاختبـارات، ولا زالـت الأبحـاث تـدور بشـأنها للتحقـق مـن صـدق الفرضيـات حولهـا. فمـع القفـزة الرقميـة في الحضـارة الإنسـانية المعاصـرة كـثرت النسـخ المحوسـبة مـن الاختبـارات. غـير أن

هـذه النسـخة الجديـدة مـن الاختبـارات أثـارت قلـق المهتمـين حـول مـدى تكافـؤ نسـخ الاختبـارات الرقميـة بنظيرتهـا التقليديـة الورقيـة. فقـد تـبزإحـدى النسـخ النسـخة الأخـرى بمميـزات تنفـرد بهـا ممـا يؤثـر علـى صـدق النتـائج. والبحـث الأول مـن هـذا العـدد يتنـاول هـذه القضيـة مـن خـلال بحـث علـمي يلقـي الضـوء علـى مـدى تكافـؤ نتـائج النسـختين.

وقـد ناقـش بحـث الدكتـورة هنـد ليـه والـذي بعنـوان «اختبـار المسـتوى الرابـع لتخصـص اللغـة العربيـة في الجامعـات الصينيـة: واقعـه والمأمـول منـه» وهـو الاختبـار الموحـد الوحيـد الـذي قـام بتنظيمـه مجلـس التوجيـه لتعليـم اللغـة العربيـة التابـع لـوزارة التربيـة والتعليـم بالصـين، والـذي يهـدف إلى قيـاس كفايـات الطلبـة اللغويـة والمهاريـة وقدرتهـم علـى التطبيـق اللغـوي المتكامـل. وتقـدم هـذه الورقـة بإيجـاز شـكل الاختبـار وآليـة تشـغيله، وتُناقـش أهميتـه وتأثيره الاجتماعـي، وكيفيـة تحسـين فاعليتـه، وتعزيـز امتثالـه لمتطلبـات القـدرة المنصـوص عليهـا في «المعيـار الوطـني لتعليـم اللغـة العربيـة في الجامعـات والمعاهـد العاليـة الصينيـة».

أمـا في ورقـة الدكتـور محمـد عبدالخالـق الـتي عنوانهـا «الاختبـارات الموضوعيـة الأكـثر اسـتخداماً والمسـتويات والنـواتج الـتي تسـتهدفها «فهـي تتنـاول بالتحليـل والتقييـم، أكـثرَ بنـود الاختبـارات الموضوعيـة شـيوعاً، والمسـتويات المعرفيـة الأكـثرَ اسـتخداماً في تلـك الاختبـارات، والنـواتج التعليميـة الـتي ينبغـي أن تحققهـا، بالإضافـة إلى العيـوب الكامنـة فيهـا، وأوجـه القصـور الـتي تعـتري مختلـف أنـواع الاختبـارات الموضوعيـة كلاً علـى حـدة.

وفي بحـث الدكتـور إسـلام يسـري والـذي وسـمه «دراسـةٌ مقارنـة بـين توجيهـاتِ المجلـس الأمريكـي لتعليـم اللغـات الأجنبيـة والإطـارِ الأوربيّ المرجعـيّ المشـترك للُّغـات» تنـاول الفـرق بـين نوعـين مـن الأطـر المرجعيـة المهمـة في تعليـم اللغـات

وتوضيـح مواضـع الاتفـاق والاختـلاف بـين توجيهـات المجلـس الأمريكـي لتعليـم اللُّغـات الأجنبيـة والإطـار الأوربي المرجعـي المشـترك للغـات، وذلـك لتوجيـه المهتمـين بمختلـف توجهاتهـم وحاجاتهـم لاختيـار الإطـار الملائـم لتبنيـه في البرنامـج.

وفي ورقـة الدكتـور عبيـدالله الفايـدي والـتي بعنـوان «تصـور مقـترح لبنـاء اختبـار معيـاري لقيـاس كفايـات معلـمي اللغـة العربيـة للناطقـين بلغـات أخـرى» عـرض فيهـا الباحـث قائمة بالكفايـات اللازمـة لمعلـمي اللغـة العربيـة لغـة ثانيـة ، وقـد حـد الكفايـات الرئيسـية بثـلاث كفايـات وهـي : الكفايـة المهنيـة والكفايـة اللغويـة والكفايـة الثقافيـة ، وانبثـق مـن هـذه الكفايـات كفايـات فرعيـة بنـاء عليهـا قـدم تصـوراً لاختبـار يقيـس هـذه الكفايـات.

وفي ورقة الدكتور عبدالعزيز أبانمي التي وسمها

«How Different Are Test-Taking Strategies with Computer-Based and Paper-Based Reading? The Case of Saudi EFL Learners»

ركـز فيهـا علـى الاسـتراتيجيات المعرفيـة الـتي يوظفهـا الطـلاب في الإجابـة علـى أسـئلة الاختبـارات والـذي سـيعود بالفائـدة علـى الباحثـين والمعلمـين علـى السـواء ، خاصـة مع تنـامي الاعتمـاد علـى التدريـس والتقييـم باسـتخدام الحاسـوب. حيـث تعـرض ورقتـه نتـائج دراسـة تستكشـف تأثـير تبايـن صيغـة عـرض المـادة المطلـوب قراءتهـا (قـراءة عـبر شاشـة الحاسـوب وقـراءة ورقيـة) علـى اسـتراتيجيات الإجابـة علـى الاختبـارات الـتي يسـتخدمها متعلمـو اللغـة الإنجليزيـة باعتبارهـا لغـةً أجنبيـة في المملكـة العربيـة السـعودية عنـد الإجابـة علـى أسـئلة الاختيـار مـن متعـدد لقيـاس الفهـم القـرائي.

ومما هـو بدهـي، فـأي منتـج ناجـح لابـد أن خلفـه فريـق عمـل جـاد، لـذا فلابـد مـن شـكر الكثيريـن مـن ذوي العلاقـة بهـذا النجـاح. ولكـن يمكـن أن نخـص هنـا سـعادة

اختبارات اللغة
Journal of Language Testing

أمـين هيئـة التحريـر د. خالـد بـن عبـدالله الصقـير الـذي لا يألـو جهـداً في الحـرص والمتابعـة ، وشـكر خـاص لسـعادة أعضـاء هيئـة تحريـر المجلـة الذيـن لا يتوانـون مشـكورين في إثـراء المجلـة بعلمهـم وعطائهـم ومقترحاتهـم. ونشـكر مجمـع الملـك سـلمان العالـمي للغـة العربيـة الـذي يتبـنى المجلـة. والشـكر موصـول للباحثـين الذيـن اختـاروا المجلـة وعـاء نشـر لمنتجاتهـم العلميـة إيمانـا بدورهـا التخصـصي المتميـز في التجسـير المعـرفي في مجـال الاختبـارات.

رئيس هيئة التحرير

أ.د. خالد بن عبدالعزيز الدامغ

مقارنة نتائج الاختبارات المحوسبة بنتائج نظيراتها الورقية في الاختبارات اللغوية

أ.د: خالد بن عبدالعزيز الدامغ[١]

أ. سارة بنت عبدالله القحطاني[٢]

مستخلص

استهدفت الدراسة الحالية مقارنة نوعين من الاختبارات، يختلفان في الوعاء الاختباري، ويفترض أن يكون بينهما تكافؤ في نتائج السمة المقيسة، إذ نظرت في مدى تكافؤ نتائج نسخ الاختبارات اللغوية الرقمية مقارنة بنظيراتها الورقية. وللوصول إلى الإجابة عن أسئلة الدراسة، فقد طُبِّقت أدواتها على مجتمع اختباري من متعلمات اللغة العربية لغةً ثانيةً في المملكة العربية السعودية، وتكونت عينة الدراسة من (٦٤) فردًا. وهدفت الدراسة في سؤالها الرئيس إلى النظر في مدى تطابق نتائج الاختبارات اللغوية المحوسبة مع نظيراتها الورقية. وأيضًا ألقت الدراسة الضوء على ما إذا كان هناك علاقة ارتباطية بين أداء مجتمع الدراسة في الاختبارات المحوسبة وخلفيتهن في استخدام الحاسب الآلي. واعتمدت الدراسة على أدوات بحثية متنوعة لتحقيق أهدافها. وخلصت إلى عدد من النتائج، كان من أبرزها بالنسبة للمحور الرئيس للدراسة أن هناك تكافؤًا بين نتائج الاختبارات المحوسبة ونتائج نظيراتها الورقية، ثم عدم وجود علاقة دالة إحصائيًا

(١) أستاذ الاختبارات في جامعة الملك سعود

(٢) باحثة دكتوراة في جامعة الملك سعود

بـين الخلفيـات المعرفيـة في الحاسـب والنتـائج في الاختبـارات اللغويـة المحوسـبة.

الكلمات المفتاحية:

اختبارات، رقمية، محوسبة، ورقية، اللغة الثانية، اللغة العربية، نتائج.

Abstract

The current study aimed to compare two types of tests, which differed in content. The study presumed that the measured attribute would show similar results. The study examined to what extent the results of the digital language tests are equivalent to their counterparts of the paper-based ones.

To find answers to the study questions, the required tools were applied to the target population of female learners of Arabic as a second language in the Kingdom of Saudi Arabia. The study sample consisted of (64) individuals.

In its core question, the study aimed to examine to what extent the results of the language computer-based tests correspond to their paper-based counterparts.

Furthermore, the study shed light on the coefficient correlation between the study population performance on the language computer-based tests and their backgrounds in using computers. It is noteworthy that the study utilized a variety of research tools to achieve its objectives.

The study showed a number of results, particularly with regard to the main theme of the study which demonstrated that the results of the language computer-based tests and the results of their paper-based counterparts were well-matched.

Eventually, the study showed that there was no statistically significant correlation between the cognitive backgrounds of the study population in using computer and their results on the language computer-based tests.

Key words:

Tests, digital, computer-based, paper-based, second language, Arabic, results.

مقدمة:

تعـد الاختبـارات إحـدى وسـائل التقييـم والتقويـم الـتي يُسـتند إليهـا في اتخـاذ القـرارات المهمـة الـتي تخـص الفـرد والمجتمـع، وتسـتخدم الاختبـارات في العديـد مـن المجـالات والعديـد مـن الأغـراض المتنوعـة، منهـا: تقييـم التحصيـل الأكاديـمي للمتعلمـين؛ وهـو مـا يعـرف بالاختبـارات التحصيليـة، وقيـاس الأداء المحتمـل لـدارس اللغـة الأجنبيـة؛ وهـو مـا يعـرف باختبـارات الاسـتعداد اللغـوي، وتحديـد المسـتوى اللغـوي للمتعلمـين؛ وهـو مـا يعـرف باختبـارات الكفايـة اللغويـة.

وقـد أوجـد التطـور التكنولـوجي في تقنيـات الحاسـب نموذجًـا آخـر للاختبـار التقليـدي، يعـرف بالاختبـار المحوسـب (CBT)، وأكـدت العديـد مـن الدراسـات مزايـا الاختبـارات المحوسـبة، ومـن هـذه الدراسـات: دراسـة طـه (2019) الـتي أثبتـت أن التغذيـة الراجعـة الـتي يوفرهـا الاختبـار المحوسـب تعـزز التعلـم، وأشـارت دراسـة Bennett (2001) إلى أن الاختبـارات المحوسـبة تقـدم الصـور والأصـوات دون الحاجـة لأجهـزة أخـرى، مـع إمكانيـة إعطـاء الدرجـة للمتعلـم فـورًا، مـا يحـد مـن الحاجـة لقيـام المصححـين بالقـراءة والمراجعـة، كمـا تسـهم عمليـة التصحيـح الإلكـتروني الـتي يتميـز بهـا الاختبـار المحوسـب في توفـير الوقـت والجهـد علـى المعلـم، واسـتثماره في مجـالات تعليميـة أخـرى .(Meijer, 2004)

وعلـى الرغـم مـن أهميـة الاختبـارات المحوسـبة ومزاياهـا المتعـددة، فـإن كثـيرًا مـن المتخصصـين يفضلـون اسـتخدام الاختبـارات التقليديـة الورقيـة، ويـرون أن الاختبـارات المحوسـبة قـد لا تتكافـأ مـع الاختبـارات الورقيـة في قيـاس التحصيـل الأكاديـمي للمتعلمـين (Davidson, 2003)، ممـا قـد يجـرح صـدق نتـائج الاختبـارات في نسـختيها الورقيـة والمحوسـبة.

وبحثت العديد من الدراسات في الاختلافات بين نتائج الاختبارات المحوسبة والورقية، ومن هذه الدراسات: دراسة حسيني وآخرين (Hosseini et al., 2013)، التي كشفت عن وجود اختلاف بين نتائج الاختبارين الورقي والمحوسب لصالح الاختبار المحوسب. ودراسة حسيني وتوروجيني (Hosseini & Toroujeni, 2017)، التي أكدت نتائجها وجود فرق بين درجات الطلبة في الاختبارين المحوسب والورقي لصالح الاختبار الورقي. ودراسة أوز وأوزتران (Öz & Özturan, 2018)، التي أظهرت نتائجها عدم وجود فرق في درجات الاختبار بين المشاركين الذين أجروا الاختبار المحوسب والذين أجروا الاختبار الورقي. ودراسة إبراهيمي وزملائه (Ebrahimi, Toroujeni and Shahbazi, 2019)، التي أكدت نتائجها عدم وجود فروق ذات دلالة إحصائية بين نتائج الطلبة في الاختبارين الورقي والمحوسب.

إلا أن النتائج التي توصلت إليها هذه الدراسات متباينة، ولم تتفق على تطابق نتائج الاختبارات المحوسبة والورقية، إذ أظهرت بعض الدراسات تطابق نتائج الاختبارات المحوسبة مع الاختبارات الورقية، على حين أظهرت نتائج الدراسات الأخرى اختلاف النتائج بين الاختبارات المحوسبة والاختبارات الورقية، ما يؤكد الحاجة إلى إجراء مزيد من الدراسات للتحقق من مدى تكافؤ نتائج الاختبارات المحوسبة مع الاختبارات الورقية، وجسر الفجوة الحاصلة بين التوجه المتسارع نحو استخدام الحاسب في التدريس من جهة، والبطء الشديد في اعتماد الاختبارات المحوسبة من جهة أخرى. كما أن أيًّا من هذه الدراسات أو غيرها -على حد علم الباحثين- لم تطبق على متعلمات اللغة العربية لغةً ثانيةً، لذا جاءت الدراسة الحالية محاوِلةً التحقق من مدى تطابق نتائج متعلمات اللغة العربية لغةً ثانيةً في الاختبارات اللغوية بنسختيها المحوسبة والورقية، مع دراسة عدد من المتغيرات التي قد تؤثر في تكافؤ النتائج.

هذه الطبعة إهـداء من المجمـع
ولا يُسمح بنشرها ورقياً أو تداولها تجارياً

مشكلة الدراسة:

ساعد انتشار الإنترنت وتقنيات الحاسب على تطوير النظم التعليمية في أنحاء متعددة من العالم، وقد شجع ذلك المعلمين على الاستفادة من هذا التطور والعمل على توظيف الحاسب في عمليات التدريس، كما أسهم هذا التطور التكنولوجي في تطور عمليات القياس والتقويم، إذ شهدت السنوات الأخيرة تطورًا في برامج تصميم الاختبارات المحوسبة. وعلى الرغم من انتشار تقنيات القياس والتقويم الحاسوبية، فإن استخدامها في الاختبارات اللغوية ليس بمستوى استخدام التقنيات الحاسوبية في التدريس (Davidson, 2003)؛ وربما يرجع التدني في استخدام الاختبارات المحوسبة والاعتماد عليها بديلًا من الاختبارات الورقية، إلى عدم ثقة بعض المتخصصين بمصداقيتها.

ونظرًا لاختلاف آراء المتخصصين نحو الاختبارات المحوسبة، وانطلاقًا من ندرة الدراسات العلمية العربية -على حد علم الباحثين- التي تؤكد تطابق نتائج الاختبارات المحوسبة مع الاختبارات الورقية من عدمه، فإن مشكلة الدراسة الحالية تتمثل في التحقق من مدى تطابق نتائج الاختبارات اللغوية المحوسبة مع نظيراتها الورقية، ودراسة بعض المتغيرات التي قد تؤثر في تكافؤ النتائج، وهي: مدى المعرفة بالحاسب الآلي؛ سعيًا إما إلى استخدام الاختبارات المحوسبة والإفادة من مزاياها المتعددة، وإما إلى الاقتناع بعدم استخدامها استنادًا إلى مسوغات علمية.

أهداف الدراسة:

تهدف هذه الدراسة إلى الآتي:

١. التحقق من مدى تطابق نتائج الاختبارات اللغوية المحوسبة مع نظيراتها الورقية، بمعرفة الفروق بين أداء المتقدمات للاختبارات في الاختبارات المحوسبة مقابل أدائهن في الاختبارات الورقية.

٢. تحديد العلاقة الارتباطية بين أداء متعلمات اللغة العربية لغةً ثانيةً في الاختبارات المحوسبة، وخلفيتهن في استخدام الحاسب الآلي.

أسئلة الدراسة:

تسعى هذه الدراسة إلى الإجابة عن الأسئلة الآتية:

١. ما مدى تطابق الاختبارات المحوسبة مع نظيراتها الورقية في نتائج قياس أداء متعلمات اللغة العربية لغةً ثانيةً في الاختبارات اللغوية؟

٢. هل توجد علاقة ارتباطية دالة إحصائيًا بين أداء متعلمات اللغة العربية لغةً ثانيةً في الاختبارات المحوسبة، وخلفيتهن في استخدام الحاسب الآلي؟

أهمية الدراسة:

تكمن أهمية هذه الدراسة في سعيها إلى التحقق من مدى تطابق نتائج الاختبارات المحوسبة مع نتائج الاختبارات الورقية في الاختبارات اللغوية، وتظهر أهمية هذه الدراسة فيما يأتي:

١. مساعدة المتخصصين على اتخاذ القرارات المناسبة في تبني الاختبارات المحوسبة من عدمه، والإسهام في جسر الفجوة الحاصلة بين الاستخدام الزائد للحاسب في العملية التعليمية من ناحية، والإحجام عن استخدام الحاسب في الاختبارات من ناحية أخرى.

٢. إثراء المكتبة العربية بأداة قياس يمكن أن تكون عونًا للباحثين في مجال تعليم اللغة العربية للناطقين بغيرها، وتتمثل في: مقياس المعرفة بالحاسب الآلي.

٣. محاولة سد ثغرة في البحث العلمي؛ وهي ندرة الدراسات العربية التي بحثت في قضية تطابق نتائج الاختبارات المحوسبة مع نظيراتها الورقية.

٤. ستوجه هذه الدراسة أنظار الباحثين والمهتمين في مجال تعليم اللغة العربية للناطقين بغيرها نحو موضوع لم يعطَ حقَّه من البحث، وعليه فإنها استفتح المجال لدراسات جديدة، وعلى مستويات تعليمية مختلفة.

حدود الدراسة:

أُجرِيت الدراسة الحالية في إطار الحدود الآتية:

١. الحدود الموضوعية: اقتصرت الدراسة على البحث في تكافؤ الاختبارات المحوسبة مع الاختبارات الورقية في قياس أداء متعلمات اللغة العربية لغةً ثانيةً في الاختبارات اللغوية، ودراسة المتغيرات التي قد تؤثر في نتائج الاختبارات المحوسبة، وهي: المعرفة بالحاسب الآلي.

٢. الحدود المكانية: طبقت الدراسة في معهد اللغويات العربية بجامعة الملك سعود، ومعهد تعليم اللغة العربية للناطقات بغيرها بجامعة الأميرة نورة بنت عبد الرحمن، في مدينة الرياض.

٣. الحدود الزمنية: طبقت هذه الدراسة على أفراد العينة في الفصل الدراسي الثاني من العام الجامعي ١٤٤٣هـ.

مصطلحات الدراسة:

يعرف الباحثان مصطلحات الدراسة إجرائيًا كما يأتي:

- الاختبارات المحوسبة (Computer-Based Tests): يُقصد بها الاختبارات التي تتم عن طريق الحاسب، إذ يعرض الاختبار على الشاشة وتجيب عنه المتعلمة باستخدام أدوات الإدخال (لوحة المفاتيح، والفأرة، وشاشة اللمس).

- الاختبارات اللغوية (Language Tests): يُقصد بها الاختبارات التي تقيس الكفاية اللغوية لمتعلمات اللغة العربية لغة ثانية.
- متعلمات اللغة العربية (Arabic Language Female Learners): متعلمات اللغة العربية لغةً ثانيةً، اللاتي يدرسن في معهد اللغويات العربية بجامعة الملك سعود، ومعهد تعليم اللغة العربية للناطقات بغيرها بجامعة الأميرة نورة بنت عبد الرحمن، بمدينة الرياض، في الفصل الدراسي الثاني من العام الجامعي ١٤٤٣هـ.

الإطار النظري والدراسات السابقة:

أولًا-مفهوم الاختبارات وخصائصها وأنواعها:

مفهوم الاختبارات:

يعرف الدامغ الاختبارات في العملية التعليمية بأنها: «أداة قياس تتضمن مثيرات أعدت بطريقة علمية بهدف تحديد حجم سمة عقلية أو صفة إنسانية» (الدامغ، ٢٠١١، ص.٤٣). أما اختبارات اللغة فيمكن تعريفها اجرائيًّا بأنها: «أداة قياس تهدف إلى تحديد مستوى المتعلمين في مهارات لغوية معينة».

خصائص الاختبارات:

تتميز الاختبارات الجيدة بخصائص عدة، منها: الصدق، والثبات، وسهولة التطبيق، والموضوعية، وغيرها. وتعد خصائص الاختبارات المختلفة مهمة في جودة الاختبارات. فمثلا سهولة التطبيق والموضوعية من مواصفات الاختبار الجيد، فالاختبار الجيد يجب أن يكون سهل التطبيق للمعلم، إذ يكون قادرًا على إعداد الاختبار وإدارته وتصحيحه في الوقت المحدد وبالإمكانات المتاحة. كما ينبغي أن تراعى الموضوعية في إعطاء النتائج، وألَّا تتأثر النتائج بذاتية المقيم.

غير أن صدق الاختبار هو السمة الأهم في الاختبارات. ويقصد بمصطلح «الصدق»: أن يقيس الاختبار فعلًا القدرة أو السمة أو الاتجاه أو الاستعداد الذي وُضع لقياسه، أي يقيس فعلًا ما يقصد أن يقيسه (إلويد، وديفيدسون، وكوم، ٢٠١٤). وتوجد العديد من الطرق للتحقق من الصدق الاختباري، وأبرزها أربع طرق رئيسة، هي (AdDamigh, 2010):

١. الصدق الظاهري: ويمكن حساب الصدق الظاهري للاختبار عن طريق التحليل المبدئي لفقراته بواسطة عدد كبير من المحكمين لتحديد إذا ما كانت هذه الفقرات تتعلق بالجانب الذي تقيسه.

٢. صدق المحتوى: ويشير إلى بيان إذا ما كانت مفردات الاختبار تمثل المجال الذي وضع الاختبار لقياسه، ودرجة مناسبة نوع المفردات لقياس ما وضعت لقياسه، ودرجة شمول عينة المفردات والطريقة التي تقاس بها محتويات مجال الاختبار.

٣. الصدق التطابقي: ويمكن حساب الصدق بهذه الطريقة بحساب مدى اتفاق درجات الأفراد على الاختبار الجديد (المراد حساب صدقه)، ودرجاتهم على اختبار آخر سبق حساب صدقه وثباته، ويقيس جوانب السلوك نفسها التي يقيسها الاختبار الجديد.

٤. صدق البناء: ويقصد به الدرجة التي يقيس فيها المقياس بناء نظريًّا أو سمة معينة دون غيرها أو مفهومًا دون غيره، ويسمى أحيانًا صدق المفهوم أو صدق التكوين الفرضي، إذ يشير إلى مدى قياس تكوين فرضي معين عن طريق التحقق التجريبي من مدى تطابق درجاته مع المفاهيم أو الافتراضات التي استند إليها الباحث في بناء المقياس.

أما الثبات فيشير إلى تجانس درجات الاختبار، أي أن يعطي الاختبار الدرجات نفسها لو أجري في وقت لاحق. ويتداخل مفهوم الثبات والصدق ويؤثر كل منهما في الآخر، فالثبات يركز على الجوانب التطبيقية في عملية القياس، على حين يركز الصدق على الجوانب النظرية (الدامغ، ٢٠١١). وهناك ثلاثة عوامل تؤثر سلبًا في ثبات نتائج الاختبارات، هي (إلويد، وديفيدسون، وكوم، ٢٠١٤):

١. **مؤثرات سببها المتعلم:**

يتعرض المتعلم المتقدم للاختبار إلى عدد من التغيرات التي قد تؤثر في ثبات نتائج الاختبارات؛ كالتعب، والأحوال النفسية، وغيرها، ما قد يؤدي إلى تغيُّر في درجة المتعلم.

٢. **مؤثرات سببها التقويم:**

قد تؤدي عملية التقويم (التصحيح) إلى تذبذب الدرجة الممنوحة نظرًا لذاتية المصحح، أو بسبب خطأ آلي، ما يؤثر في ثبات نتائج الاختبار.

٣. **مؤثرات سببها إدارة الاختبار:**

قد يؤثر عدم ثبات أسلوب الإدارة، مثل: اختلاف العوامل البيئية، والعوامل المتعلقة بالطباعة، وغيرها من العوامل المتعلقة بإدارة الاختبار، في ثبات نتائج الاختبارات.

وذكر النبهان (٢٠٠٤) أن معامل ثبات أداة القياس يتأثر بعدد من العوامل، ومنها: طول الاختبار، إذ تعطي الاختبارات الأكثر طولًا درجات أكثر ثباتًا. وتجانس مجموعة المفحوصين، إذ يزداد معامل ثبات الاختبار إذا ازدادت درجة تباين مجموعة المفحوصين، فإذا طبق الاختبار على مجموعة من المتفوقين فقط أو ضعيفي التحصيل فقط فإن معامل الثبات سيكون منخفضًا، ما يشير إلى أهمية تطوير أداة قياس بعد تطبيقها على عينة ممثلة للمجتمع المراد تعميم نتائج الاختبار عليه، وذلك بأن تغطي العينة كلَّ طبقات المجتمع ومستوياته. وبما

أن معامل الثبات يعتمد على درجة التباين في درجات المفحوصين، فإنه يزداد تبعًا للعوامل التي تحسن مستوى ذلك التباين، وذلك بالأداء على الفقرات التي تتمتع بمعاملات صعوبة متوسطة، لأن الفقرات السهلة جدًّا أو الصعبة جدًّا لا تظهر الفروق الفردية بين المفحوصين، ما يجعل معامل الثبات منخفضًا. كما أن الفقرات التي تصحح بموضوعية عالية تؤدي إلى تحسين وزيادة معامل الثبات.

وتختلف الطرق المستخدمة لقياس الثبات، ومن أكثرها شيوعًا: ثبات الاستقرار Stability Reliability، وثبات التكافؤ Equivalence Reliability، وثبات الاتساق الداخلي Internal Consistency. ويعرف مقياس ثبات الاستقرار بطريقة الاختبار وإعادة الاختبار Test–Retest–Method، إذ يطبق الاختبار على مجموعة من الأشخاص، ثم يعاد تطبيق الاختبار على المجموعة نفسها في وقت لاحق، ويتبع ذلك حساب معامل الارتباط بين درجات أفراد المجموعة على ذلك الاختبار في المرتين (النبهان، ٢٠٠٤). أما ثبات التكافؤ فيحسب بإعطاء شكلين متكافئين لاختبار معين للمجموعة نفسها وفي الوقت نفسه (بعد فاصل زمني قصير)، وحساب معامل الارتباط بينهما (النبهان، ٢٠٠٤). ويتطلب حساب ثبات الاستقرار وثبات التكافؤ الحصول على بيانات من جلستين اختباريتين، على حين يتطلب حساب ثبات الاتساق الداخلي للاختبار حساب مدى ترابط الاستجابات على الفقرة الواحدة مع درجة الاختبار الكلية (النبهان، ٢٠٠٤).

أنواع الاختبارات:

تتنوع اختبارات اللغة، وتختلف باختلاف الهدف من تطبيقها، ومن أبرزها: اختبارات التصنيف، واختبارات الكفاية اللغوية. وتهدف اختبارات التصنيف إلى قياس المقدرة اللغوية لدى المتعلمين، حتى يُصنَّفوا في المستوى المناسب لقدراتهم اللغوية، بما يحقق التجانس بين أفراد المجموعة (إلويد، وديفيدسون، وكوم، ٢٠١٤)، وبما يضمن توزيع المتعلمين الجدد كل حسب مستواه في مجموعة من

المجموعات التي تناسبه (الدامغ، ٢٠١١). وقد طبقت الدراسة الحالية اختبار التصنيف بمعهد اللغويات العربية بجامعة الملك سعود على عينة الدراسة، لضمان تجانس مستوياتهم اللغوية.

أما اختبارات الكفاية اللغوية فتهدف إلى قياس المقدرة اللغوية الكلية للمتعلمين، ما يفيد في التعرف إلى مستوى مقدرة المتقدم للاختبار في مهارة لغوية معينة (إلويد، وديفيدسون، وكوم، ٢٠١٤).

ثانيًا-الاختبارات المحوسبة:

تطورت أساليب القياس والتقويم نتيجة للتقدم والتطور التكنولوجي الذي يشهده العالم في الوقت الحاضر، ولعل أبرز التطورات التكنولوجية ظهور الاختبارات المحوسبة (الرقمية)، واتجاه كثير من المؤسسات التعليمية لاستخدامها في تقويم المتعلمين. وفيما يأتي عرض لمفهوم الاختبارات المحوسبة، ونشأتها، وأنواعها، وأبرز إيجابياتها، وسلبياتها، وفاعليتها في قياس التحصيل الدراسي للمتعلمين.

مفهوم الاختبارات المحوسبة:

تعد الاختبارات المحوسبة إحدى الوسائل الدقيقة في تقويم أداء الطلبة إلكترونيًّا، وتعرف بأنها: «نسخة مطابقة للاختبار الورقي من حيث الشكل والمحتوى وتسلسل الفقرات، والفرق بين الاختبار المحوسب والاختبار الورقي، أن الأول يقدم عن طريق الحاسب، على حين يقدم الاختبار الثاني باستخدام الطرق الاعتيادية على شكل ورقي» (Bookman, 1989, 26).

نشأة الاختبارات المحوسبة:

بدأ استخدام الاختبارات المحوسبة في أواخر الستينيات من القرن الماضي، تحديدًا في وزارة الدفاع الأمريكية، إذ وجهت الوزارة الباحثين في علم النفس الإكلينيكي

نحو إعداد الاختبارات المحوسبة، ثم بعد ذلك بدأ استخدام الاختبارات المحوسبة في مجال التعليم في أواخر الثمانينيات من القرن الماضي (Russell et al., 2003).

أنواع الاختبارات المحوسبة:

ثَمَّ أنواع مختلفة للاختبارات المحوسبة، ومنها: الاختبارات الخطية المحوسبة، والاختبارات التكيفية المحوسبة، وتختلف هذه الأنواع في طريقة التطبيق؛ فالاختبارات الخطية المحوسبة تُقدم فيها جميع فقرات الاختبار لجميع الطلبة بالتساوي، أما في النوع الآخر الذي يسمى الاختبارات التكيفية المحوسبة، فتُعرض بعض الفقرات على المتعلمين لتحديد قدراتهم، وبناء على هذه القدرة تُقدَّم الفقرات اللاحقة بما يناسب مستوى المتعلم (Cisar, Radosav, Markoski, Pinter & Cisar, 2010).

وتتعدد برامج تصميم الاختبارات المحوسبة، فمنها ما هو مجاني، ومنها ما هو برسوم، ومن الأمثلة على البرامج المجانية: برنامج (Question Tools)، وبرنامج (Quiz Faber)، وبرنامج (Qedoc Quiz Maker)، إضافة إلى مواقع الاختبارات الإلكترونية على شبكة الإنترنت، مثل: موقع (eQuizzer)، وموقع (Quiz School)، وتوجد على هذه المواقع العديد من القوالب الجاهزة التي تيسر عملية تصميم الاختبارات. أما البرامج غير المجانية فهي كثيرة ومتنوعة، منها على سبيل المثال لا الحصر: برنامج (Book widgets)، وبرنامج (Quiz Builder)، وبرنامج (Quiz Maker).

إيجابيات الاختبارات المحوسبة:

أشار عدد من الباحثين إلى إيجابيات الاختبارات المحوسبة، ومن هذه الإيجابيات ما ذكره ريان وزملاؤه (Ryan, Scott, Freeman & Patel, 2000): التفاعلية، والمرونة، وتوفير الوقت، والتغذية الراجعة الفورية، واختصار الموارد، والاحتفاظ بالسجلات. ويضيف بينيت Bennett (٢٠٠١) مزايا أخرى هي: توفير تكلفة

الطباعة، وسهولة مراجعة بنود الاختبار، وتقديم مثيرات حركية بالصوت والصورة والرسوم المتحركة دون الحاجة لأجهزة أخرى، وإعطاء الدرجة للطلبة فورًا ودون تحيز، والحد من الحاجة لقيام المصححين بالقراءة والمراجعة.

ويرى كولز وكوك وبليك (Coles, Cook & Blake, 2007) أن استخدام الاختبارات المحوسبة يسهل عملية جمع البيانات، وعملية التصحيح، ووضع التعليمات الخاصة بعملية الاختبار، إضافة إلى توفير الكلفة الاقتصادية، والجهد المبذول عند تقديم الاختبار؛ من حيث الطباعة، والتطبيق والملاحظة، ونقل الدرجات، وتوفير عدد أقل من القائمين على تنظيم ومراقبة الاختبارات.

وبناء على ما سبق يمكن تصنيف هذه المزايا إلى: مزايا متعلقة بتوفير الوقت والجهد والتكلفة في تصحيح الاختبار واستخراج النتائج، ومزايا تتعلق بإمكانية تصميم الاختبارات بسهولة ودون الحاجة لوجود أجهزة أخرى في أثناء الاختبار كما هو الحال في الاختبارات الورقية، فيمكن على سبيل المثال: إضافة التسجيلات الصوتية في اختبارات الاستماع المحوسبة بسهولة، وإضافة الصور والرسوم المتحركة ومقاطع الفيديو في اختبار المفردات، ونحو ذلك.

وقد أجمل الدامغ وعبد الخالق (٢٠١٢) إيجابيات الاختبارات المحوسبة مقارنة بالاختبارات الورقية فيما يأتي:

١. ارتفاع درجة صدق أدوات القياس وثباتها.
٢. قلة الحاجة لعمليات الطباعة والتصوير.
٣. سرعة وسهولة إجراء الاختبارات.
٤. الوقت المستغرق للأداء أقصر.
٥. الدقة المتناهية في التقويم ورصد النتائج.

٦. سرعة الحصول على النتائج.

٧. موضوعية التقويم وعدالته.

٨. قلة عدد العاملين؛ فليست هناك حاجة لإدخال الدرجات يدويًّا.

٩. إمكان تقديم صوت، وصور متحركة مصاحبة للأسئلة.

ومن المزايا المهمة في الاختبارات المحوسبة التي تجدر الإشارة إليها: زيادة معاملات الصدق والثبات، إذ يتيح هذا النوع من الاختبارات إمكانية استرجاع معلومات من الحاسب عن الوقت المستغرق للإجابة عن كل سؤال، والأسئلة التي غيّر المتعلم إجابته فيها، والأسئلة التي تجاوزها المتعلم وعاد ليجيب عنها، وهذه الخاصية يمكن استخدامها في تحليل معامل صدق الأداة. كما أن نظام التصحيح الآلي يزيد من درجة ثبات التقويم، وذلك بالابتعاد عن ذاتية المقوِّم التي قد تتأثر بعوامل مختلفة (الدامغ، وعبد الخالق، ٢٠١٢).

سلبيات الاختبارات المحوسبة:

تتضمن أدوات القياس والتقويم المحوسبة عددًا من السلبيات التي قد تثير الجدل حول صدقها وفاعليتها وتحد من انتشارها، ومن هذه السلبيات: قضية إعداد الاختبارات المحوسبة وأنها قد تتطلب مهارات عالية في التعامل مع التقنيات المستخدمة في تصميم الاختبارات المحوسبة، ما يستوجب تدريب المعلمين على إنشائها (Bennett, 2001)، وأن إجراء الاختبارات المحوسبة يتطلب توفير عدد كافٍ من أجهزة الحاسب، وصيانتها، كما أن انقطاع الإنترنت قد يعوق تنفيذ الاختبارات المحوسبة (Friedrich & Julius, 2009; Marks & Cronje, 2008)، وكذلك قد لا يستطيع المتقدم للاختبار رؤية جميع جوانب الاختبار، أو طرح الاستفسارات عندما يعوقه فهم الأسئلة (الدامغ، وعبد الخالق، ٢٠١٢). إضافة إلى محدودية الخيارات في كتابة الاختبار واختيار الأسئلة؛ حيث يجد المختبر نفسه مقيدًا بأسئلة قابلة للتصحيح

الآلي؛ لذا تقتصر الاختبارات المحوسبة في غالبها على الأسئلة الموضوعية، ولا تزال البرمجيات في تصحيح الأسئلة المقالية محدودة.

فاعلية الاختبارات المحوسبة:

يشير عدد من الدراسات إلى فاعلية الاختبارات المحوسبة وأهميتها في قياس التحصيل الدراسي للمتعلمين، ومن هذه الدراسات: دراسة ودفيلد Woodfield (٢٠٠٣)، التي أظهرت نتائجها أن الاختبارات الورقية لا تقيس بوضوح النمو الدراسي للمتعلمين، ولا تقدم رؤية للمعلم عن كيفية تعديل المنهج الدراسي سعيًا لتحسين مخرجات التعلم. وتؤكد دراسة أخرى أن الاختبارات المحوسبة أكثر فاعلية من التقليدية، لأنها تساعد المعلم على تقويم درجة استيعاب المتعلمين، وذلك لأن استجاباتهم للاختبارات وانتقالاتهم بين الأسئلة وتعديلاتهم مسجلة إلكترونيًا (Kearsley, 2000)، ما يساعد المعلم على معرفة مواطن الصعوبة لدى المتعلمين ومعالجتها.

ثالثًا-دراسات قارنت بين نتائج الاختبارات المحوسبة ونظيراتها الورقية:

أجريت العديد من الدراسات الأجنبية التي هدفت إلى مقارنة الاختبارات المحوسبة بالاختبارات الورقية بهدف التحقق من تطابق نتائج الاختبارات المحوسبة مع نتائج الاختبارات الورقية في قياس التحصيل الدراسي في مراحل دراسية مختلفة، أو في قياس الكفاية اللغوية لمتعلمي اللغة الثانية، وقد طبقت هذه الدراسات في أمريكا، وأوروبا، ودول أخرى متقدمة. أما الدراسات العربية في هذا المجال فتعد دراسات نادرة جدًّا، ولم يقف الباحثان على دراسات عربية بحثت في تكافؤ نتائج الاختبارات المحوسبة والورقية إلا دراستين فقط، وذلك حسب الجهد والاطلاع، كما أن هاتين الدراستين لم تطبقا على متعلمي اللغة الثانية، ولم تبحثا في تكافؤ الاختبارات المحوسبة مع نظيراتها الورقية في الاختبارات اللغوية.

ومـن أوائـل الدراسـات الـتي هدفـت إلى التعـرف إلى درجـة التكافـؤ بـين الاختبـارات الورقيـة ونظيراتهـا المحوسـبة في المرحلـة الجامعيـة: دراسـة نيومـان وبايـدن (Neuman & Baydoun, 1998)، الـتي شـارك فيهـا (٤١١) طالبًـا وطالبـة، وقـد أجريـت عليهـم عشـرة اختبـارات قبليـة وبعديـة. وتوصلـت الدراسـة إلى عـدم وجـود فـروق ذات دلالـة إحصائيـة بـين الاختبـارات الورقيـة والمحوسـبة. ودراسـة ديبـير وفاسـير (Debeer & Visser, 1998) الـتي هدفـت إلى مقارنـة نتـائج الاختبـارات الورقيـة ونظيراتهـا المحوسـبة بقصـد التعـرف إلى درجـة التكافـؤ بـين الاختبارين في قيـاس التحصيـل الـدراسي. وأجريـت الدراسـة على (٦١٣) طالبًـا وطالبـة في المرحلـة الثانويـة في جنـوب إفريقيـا. وخلصـت نتائجهـا إلى أن الاختبـار الـورقي كان الأعلـى في النتـائج.

ومـن الدراسـات الـتي هدفـت إلى معرفـة الفـرق بـين الاختبـارات الورقيـة والمحوسـبة عـن طريـق دراسـة العوامـل الـتي قـد تؤثـر في تكافـؤ النتـائج: دراسـة كلارينـا ووالاس (Clariana & Wallace, 2000)، إذ بحثـت الدراسـة في أربعـة عوامـل، هي: الجنـس، والمعرفـة بالمحتـوى، والمعرفـة بالحاسـب، والتنافسـية. وطبقـت الدراسـة على (١٠٥) مـن الطـلاب والطالبـات في المرحلـة الجامعيـة. وقـد تعـرض أفـراد العينـة لنسـختين مـن الاختبـار نفسـه، ولكـن إحداهمـا ورقيـة والأخـرى محوسـبة، وأظهـرت نتـائج الدراسـة تفوقًـا للطلبـة الذيـن تعرضـوا للاختبـار المحوسـب، ولـم توجـد فـروق في العوامـل المتعلقـة بالجنـس والتنافسـية والمعرفـة بالحاسـب، على حـين وُجِـد فـرق في عامـل المعرفـة بالمحتـوى.

أمـا دراسـة سـايم وهورتـون فقـد هدفـت إلى معرفـة الفـروق في أداء الطلبـة بـين الاختبـار الـورقي والمحوسـب، ومعرفـة اتجاهـات الطلبـة نحـو الاختبـار المحوسـب. وطبقـت على (٢٠) طالبًـا وطالبـة في الصـف الثالـث الابتـدائي في المملكـة المتحـدة. وخلصـت نتائجهـا إلى أن (٥٠٪) مـن الطـلاب كان أداؤهـم أفضـل في الاختبـار الـورقي، علـى حـين كان أداء (٢٥٪) منهـم أفضـل في الاختبـار المحوسـب، وكان أداء (٢٥٪)

متماثلًا في الاختبارين الورقي والمحوسب، كما أظهرت نتائج الدراسة أن الطلاب يفضلون الاختبارات المحوسبة على الاختبارات الورقية (Sim & Horoton, 2005).

وفي دراسة آل ثاني (٢٠٠٨) التي هدفت إلى البحث عن مدى التكافؤ بين نتائج اختبار رافن بنسختيه المحوسبة والورقية، لدى عينة من المتعلمين والمتعلمات بجامعة قطر، والكشف عن العوامل المؤثرة في تكافؤ نتائج النسختين، مثل: قلق الاختبار، والجنس، وقلق الحاسب، والاتجاه نحو الحاسب. توصلت الباحثة إلى عدم وجود فروق ذات دلالة إحصائية بين درجات المتقدمين والمتقدمات للاختبار الورقي والمحوسب تبعًا لاختلاف مستويات قلق الاختبار وقلق الحاسب والاتجاه نحو الحاسب.

وبالمثل، فقد بحث الخزي وزكي (٢٠١١) مدى التكافؤ بين الاختبارات المحوسبة والاختبارات الورقية في قياس التحصيل الدراسي، ومدى تأثير تعرض الطلبة للاختبارات المحوسبة في اتجاهاتهم نحوها. واعتمدت الدراسة المنهج التجريبي، وطُبِّقت على (٣١٦) طالبًا وطالبة في كلية التربية بجامعة الكويت. وأظهرت نتائج الدراسة تكافؤ الاختبارات المحوسبة والورقية في قياس التحصيل الدراسي للطلبة مع وجود فروق ذات دلالة إحصائية في الوقت اللازم لأداء الاختبار لصالح الاختبارات المحوسبة، كما أظهرت النتائج ارتفاع اتجاهات الطلبة نحو الاختبارات المحوسبة بسبب تعرضهم لها. وأوصت الدراسة بإجراء المزيد من الدراسات في التخصصات الأخرى.

ومن الدراسات التي عنيت بالعلاقة بين المعرفة بالحاسب ونتائج الاختبارات المحوسبة: دراسة حسيني وآخرين (Hosseini et al., 2013)، إذ هدفت الدراسة إلى البحث فيما إذا كان هناك أي فرق بين نتائج الاختبار المستند إلى الحاسوب المتواضع عليها اختصارًا بـCBT، ونتائج الاختبار الورقي المتواضع عليها اختصارًا بـPBT، لقياس الاستيعاب في أثناء قراءة النصوص، والنظر في العلاقة بين خبرة الطلاب في التعامل

مع الحاسب الآلي ودورها في أداء الاختبار المحوسب CBT. وقد طبقت الدراسة على (١٦٢) طالبًا، إذ أُجرِي اختباران متكافئان أحدهما ورقي والآخر محوسب لمجموعة واحدة من متعلمي اللغة الإنجليزية لغة أجنبية EFL في زمنين مختلفين. وأظهرت نتائج الدراسة وجود اختلاف بين نتائج الاختبارين الورقي والمحوسب لصالح الاختبار الورقي، كما أظهرت النتائج أن خبرة الطلاب السابقة حول استعمال الحاسب لم يكن لها تأثير كبير في أداء الطلاب في الاختبار المحوسب.

وفي دراسة أخرى هدفت إلى التحقق من صدق نتائج الاختبار المحوسب، وذلك عن طريق مقارنة نتائجه بنتائج الاختبار الورقي. وطبقت على (١٠٠) طالب وطالبة من طلبة الدراسات العليا بجامعة شاباهار الساحلية. أظهرت نتائج الدراسة تفوق نتائج الاختبار المحوسب CBT على نظيره PPT، وأشارت النتائج إلى أن المعرفة السابقة بالحاسب، وتوجه الطلاب نحوه، لم يكن لهما تأثير كبير في تميز درجات الطلاب الذين أجري عليهم الاختبار المحوسب CBT، كما تبين أن نسبة كبيرة من المشاركين قد أبدوا اهتمامًا وميولًا للاختبار المحوسب على نظيره الورقي، ما يشير إلى أهمية هذا النوع من الاختبارات بديلًا من الاختبار التقليدي (الورقي) (Khoshsima, Morteza & Toroujeni, 2017).

وبحثت دراسة خوشيما وآخرين (Khoshsima et al., 2017) مقارنة نتائج الطلاب في الاختبارات المحوسبة بنتائجهم في الاختبارات الورقية، للتحقق من تطابق نتائج الاختبارات المحوسبة والورقية. وطبقت الدراسة على (٣٠) طالبًا وطالبة ممن يتعلمون اللغة الإنجليزية لغة أجنبية في المرحلة الجامعية، وذلك عن طريق مقارنة أدائهم في اختبار المفردات، وقد صمم الباحثون اختبارين متماثلين واحدًا بنسخة ورقية والآخر بنسخة رقمية. وأظهرت النتائج وجود علاقة وثيقة بين ميول الطلبة والنتائج التي حصلوا عليها، كما أظهرت نتائج الدراسة عدم وجود اختلاف في النتائج يعزى إلى الجنس، أي أن فئة المشارك لم تؤثر في الأداء، وبناء

على ذلك خلصت الدراسة إلى أنه يمكن عدُّ النسخة المحوسبة من الاختبار بديلًا مناسبًا لطلاب المرحلة الجامعية في إيران.

وفي دراسة أخرى، قام حسيني وتوروجيني (Hosseini & Toroujeni, 2017) بمقارنة درجات الاختبار الورقي PBT ونظيره المحوسب CBT في قياس تحصيل مهارة اللغة الإنجليزية في جامعة بايمي نور (Payame Noor PNU) لطلاب السنة الأولى بالجامعة، وطبقت الدراسة على (٦٠) طالبًا وطالبة. وأظهرت النتائج أن أداء الطلاب في الاختبار الورقي PBT أفضل من أدائهم في الاختبار المحوسب CBT، كما أن معظم المشاركين أبدوا ميولًا عالية للاختبار المحوسب CBT مقارنة بنظيره الورقي PBT، ولكن نظرًا لوجود بعض المسوغات التي أبدوها واعتيادهم إجراءَ الاختبارات تقليديًا، كان أداؤهم أفضل في اختبار الورقي المصطلح عليه بـPBT. ومن أبرز التوصيات التي خرجت بها هذه الدراسة ضرورة إجراء المزيد من الدراسات المقارنة في سياقات التعليم العالي قبل اعتماد الاختبار المحوسب CBT بديلًا من الورقي PBT أو إدراجه في النظام التعليمي.

كما هدفت دراسة حسيني وتوروجيني (Hosseini & Toroujeni, 2017) إلى مقارنة الاختبار الورقي PBT ونظيره المحوسب CBT في قياس تحصيل طلبة جامعة بايمي نور (Payame Noor (PNU في مهارة اللغة الإنجليزية، وذلك للتحقق من صدق الاختبارات المحوسبة، كما هدفت الدراسة إلى البحث في العلاقة بين الفئة الجنسية للمشاركين ومدى إلمامهم ومعرفتهم المسبقة بالحاسوب، وفي أدائهم في الاختبار المحوسب CBT، وقد طبقت الدراسة على (١٠٠) طالب وطالبة. وأظهرت النتائج وجود فرق بين درجات الطلبة في الاختبارين المحوسب والورقي لصالح الاختبار الورقي PPT، وقد تكون نتيجة انخفاض الدرجات في الاختبار المحوسب CBT بسبب عدم التهيئة الجيدة للاختبار المحوسب بحسب ما ذكره المشاركون في المقابلة والنقاش الجماعي، كما أظهرت النتائج أنه لا يوجد تأثير لمتغير الجنس

في نتائج الاختبار المحوسب CBT، كما أنه لا يوجد أي ارتباط بين الإلمام بالحاسب ونتيجة الأداء في الاختبار المحوسب.

وفي دراسة حديثة، هدفت إلى التحقق من موثوقية ودقة الاختبار المحوسب مقارنة بنظيره الورقي، طبق الباحثان دراستهما على (٩٧) طالبًا جامعيًّا ملتحقين ببرنامج اللغة الإنجليزية لغة أجنبية (EFL) المعد خصوصًا لمعلمي اللغة الإنجليزية. ووُزِّع المشاركون عشوائيًّا إلى مجموعة تجريبية أجرت اختبارًا يعتمد على الحاسب، ومجموعة أخرى قياسية أُجري عليها اختبار مماثل لكنه وفق النظام التقليدي الورقي. وقد أشارت النتائج إلى أن وضع تهيئة الاختبار لم يكن له أي تأثير في موثوقية وصلاحية الاختبارات التي تدار في أي من الاتجاهين، كما أظهرت النتائج أنه لا يوجد فرق كبير في درجات الاختبار بين المشاركين الذين أجروا الاختبار المستند إلى الحاسوب والذين أجروا الاختبار الورقي. ونوقشت النتائج من حيث فكرة إمكانية دمج تكنولوجيا الحاسب في المناهج الدراسية ليس لأغراض التدريس فقط وإنما أيضًا لأغراض القياس والتقويم (Öz & Özturan, 2018).

وهدفت دراسة إبراهيمي وزملائه (Ebrahimi, Toroujeni and Shahbazi, 2019) إلى مقارنة متوسط الدرجات التي حصل عليها أفراد العينة في الاختبار المحوسب CBT والاختبار الورقي PBT، لمعرفة إذا ما كان أداء المتقدمين للاختبار يتأثر بكيفية وضع الاختبار وتهيئته، لذلك أُجريت هذه الدراسة للبحث في مدى اختلاف الدرجات بين الاختبار الورقي والاختبار المحوسب، والبحث أيضًا في العلاقة بين فئة الجنس، وتفضيل نمط الاختبار على أداء المتقدمين للاختبار المحوسب، وطبقت الدراسة على (١٢٠) طالبًا وطالبة باستخدام أداتي الاختبار والمقابلة. وقد أشارت النتائج إلى عدم وجود فروق ذات دلالة إحصائية بين نتائج الطلبة في الاختبارين الورقي والمحوسب؛ أي إنه لا يوجد اختلاف بين الاختبار الورقي والمحوسب، ولا توجد علاقة ذات دلالة إحصائية بين المتغيرات التي ذكرت سابقًا (الجنس، والتفضيل) وبين نتائج أداء الاختبار المحوسب CBT.

وبالمثل، فقد بحثت دراسة ياو (Yao, 2020) مقارنة أداء المتقدمين للاختبار المحوسب CBT مع نظيره الورقي PBT وفقًا للمعيار المرجعي الأوروبي المشترك للغات CEFR المختلفة. وقد أجريت الدراسة على مجموعة من الطلاب وعددهم (٢٨٩) طالبًا من مختلف المدارس الثانوية في ماكاو، وقد أجري اختبار تجريبي للغة الإنجليزية الأكاديمية (TAE) في إحدى الجامعات المحلية. وأشارت النتائج إلى عدم وجود اختلاف بين أداء المتقدمين للاختبار الورقي والمحوسب، وخصوصًا فيما يتعلق بنمطية الاختبار عبر المستويات الستة CEFR المختلفة، ما عدا في المستوى الثاني CEFR A٢ فقد وجد فرق إحصائي لصالح الاختبار المحوسب CBT. وقد أوصت الدراسة بضرورة تحويل نمط الاختبارات الجامعية من اختبارات تقليدية ورقية PBT إلى اختبارات محوسبة CBT.

وفي دراسة أخرى، قام همهيس وجلاس وميلسين (Hamhuis, Glas and Meelissen, 2020) باستكشاف الاختلافات المحتملة في الأداء بين الاختبار الورقي PBT وتقويم الأجهزة الذكية في الرياضيات والعلوم بين طلاب المدارس الابتدائية الهولندية. ففي ربيع عام ٢٠١٧م أجري الاختبار الورقي PBT واختبار الأجهزة الذكية (اللوحية) لـ(٥٣٢) طالبًا هولنديًّا في الصف الرابع. وخلصت الدراسة إلى أنه لا توجد اختلافات كبيرة في مقاييس قدرة الطالب بين الورق واختبارات الأجهزة الذكية (اللوحية) للرياضيات والعلوم، كما أظهرت النتائج وجود فرق ضئيل يعزى لمتغير الجنس، إذ تفوقت الإناث على الذكور في اختبار الأجهزة الذكية (اللوحية)، على الرغم من عدم وجود فروق بين الجنسين في نتائج الاختبار الورقي PBT.

ويخلص الباحثان من استعراض الدراسات السابقة إلى ما يأتي:

١. حسب الجهد والاطلاع، فلا توجد دراسات عربية في مجال تعليم اللغة العربية لغير الناطقين بها، تناولت موضوع التحقق من تطابق نتائج الاختبارات المحوسبة مقارنة بنظيراتها الورقية في قياس الكفاءة اللغوية لغير الناطقين باللغة العربية، وستكون هذه الدراسة هي الأولى التي تبحث في تطابق نتائج الاختبارات المحوسبة مع نظيراتها الورقية في قياس الكفاءة اللغوية لمتعلمي اللغة العربية لغةً ثانيةً.

٢. وقف الباحثان على دراستين سابقتين باللغة العربية طبقت على متعلمين عرب في اختبار رافن، وفي مقرر الحاسب، وهما: دراسة آل ثاني (٢٠٠٨)، ودراسة الخزي وزكي (٢٠١١). أما الدراسات الأجنبية فقد أجريت في أنحاء مختلفة من العالم، وطبقت على متعلمين من ذوي خلفيات لغوية متنوعة، واستخدمت الاختبارات اللغوية وغير اللغوية أداة بحثية.

٣. لم تتفق نتائج الدراسات السابقة على تطابق نتائج الاختبارات المحوسبة مع نتائج الاختبارات الورقية، إذ أظهرت بعض الدراسات تطابق نتائج الاختبارات المحوسبة مقارنة بالاختبارات الورقية، على حين أظهرت نتائج الدراسات الأخرى اختلاف النتائج بين الاختبارات المحوسبة والاختبارات الورقية، ما يؤكد الحاجة إلى إجراء مزيد من الدراسات للتحقق من مدى تكافؤ نتائج الاختبارات المحوسبة مع الاختبارات الورقية أو تباينها.

٤. استفاد الباحثان كثيرًا من هذه الدراسات في تكييف أدوات قياس تلائم اللغة العربية، وتفيد الباحثين مستقبلًا في تطوير دراسات تهتم بتجويد الاختبارات المحوسبة لغير الناطقين بالعربية.

٥. مع أن الدراسة الحالية اعتمدت اعتمادًا كبيرًا على هذه الأفكار والدراسات السابقة، واستفادت منها في الإطار النظري، ومنهجية البحث، لكنها اختلفت عنها في أن هذه الدراسة طبقت في بيئة مختلفة، وطبقت على متعلمات اللغة العربية غير الناطقات بها، وهي دراسة غير مسبوقة حسب الجهد والاطلاع، ولذلك فإن نتائج هذه الدراسة إما أن تعزِّز النتائج السابقة التي أكدت تطابق نتائج الاختبارات المحوسبة مع الاختبارات الورقية، وإما أن تنفيها. كما اتخذت هذه الدراسة عددًا من الإجراءات لضمان تجانس أفراد العينة، وإبعاد العوامل التي قد تؤثر في ثبات نتائج الاختبار، وهو ما يعرف بمصطلح (Homogeneous Sampling) (Deviant, 2011).

٦. تتفق الدراسة الحالية مع الدراسات السابقة في بحث العوامل التي قد تؤثر في نتائج الاختبارات مثل: المعرفة بالحاسب.

اولًا- منهج الدراسة:

اتبعت الدراسة الحالية المنهج الكمي في جمع بياناتها وتحليلها، إذ اعتمدت المنهج التجريبي في دراسة مدى تطابق الاختبارات المحوسبة مع نظيراتها الورقية في قياس أداء متعلمات اللغة العربية لغةً ثانيةً في الاختبارات اللغوية.

ثانيًا- مجتمع الدراسة:

مجتمع هذه الدراسة هو متعلمات اللغة العربية لغةً ثانيةً ممن يدرسن في المستوى الثالث والرابع والماجستير في معهد اللغويات العربية بجامعة الملك سعود، والمستوى الثالث والرابع بمعهد اللغة العربية للناطقات بغيرها بجامعة الأميرة نورة بنت عبد الرحمن، بمدينة الرياض.

ثالثًا- عينة الدراسة:

تتمثل عينة الدراسة في جميع أفراد مجتمع الدراسة، وهنَّ متعلمات اللغة العربية لغةً ثانيةً في المستوى الثالث والرابع والماجستير في معهد اللغويات العربية بجامعة الملك سعود، والمستوى الثالث والرابع بمعهد اللغة العربية للناطقات بغيرها بجامعة الأميرة نورة بنت عبد الرحمن، ويبلغ عددهن (٦٤) متعلمة. وقد اختيرت عينة الدراسة بأسلوب الحصر الشامل.

رابعًا- خصائص العينة:

بلغ عدد مفردات عينة الدراسة (٦٤) متعلمة للعربية لغةً ثانيةً، يمثلون خمسًا وعشرين دولة حول العالم، وهذا التنوع في الجنسيات والثقافات أعطى الدراسة بُعدًا إضافيًّا مهمًّا انعكس على نتائجها. وتصنف عينة الدراسة حسب اختبار التصنيف الذي أجراه الباحثان ضمن المستوى الثالث والرابع والمجتاز. وقد أجريت الدراسة على عينة من الإناث، لإبعاد تأثير متغير الجنس في نتائج الاختبارات. وفيما يأتي وصف لهذه العينة تبعًا للبيانات الأولية.

١. **المستوى الدراسي:**

جدول رقم (١) توزيع العينة وفقًا للمستوى الدراسي

المستوى الدراسي	العدد	النسبة المئوية
الثالث	٤٠	٦٢,٥٠٪
الرابع	٢٠	٣١,٢٥٪
الماجستير	٤	٦,٢٥٪
المجموع	٦٤	١٠٠٪

٢. **العمر:**

جدول رقم (٢) توزيع العينة وفقًا للعمر

العمر	العدد	النسبة المئوية
٢٠ سنة فأقل	١٢	٪١٨,٧٥
٢١–٢٤ سنة	٣٠	٪٤٦,٨٨
٢٥ سنة فأكثر	٢٢	٪٣٤,٣٧
المجموع	٦٤	٪١٠٠

٣. **الجنسية:**

جدول رقم (٣) توزيع العينة وفقًا للجنسية

الجنسية	العدد	النسبة
كازاخية	١	١,٥٦
نيجيرية	٦	٨,٣٩
تشادية	١	١,٥٦
باكستانية	١	١,٥٦
عاجية	٣	٩,٦٤
كوسوفية	٤	٦,٢٥
ألبانية	١	١,٥٦
توغولية	١	١,٥٦
سنغالية	١	١,٥٦
صومالية	١	١,٥٦
رواندية	٢	٣,١٣
بريطانية	١	١,٥٦

الجنسية	العدد	النسبة
مالية	١	١,٥٦
بوركينية	١	١,٥٦
تايلاندية	٢	٣,١٣
نيبالية	١	١,٥٦
بنينية	٣	٩,٦٤
ماليزية	١	١,٥٦
هندية	٥	٧,٨١
سريلانكية	٢	٣,١
فلبينية	١٣	٢٠,٣١
غينية	٣	٩,٦٤
هولندية	٢	٣,١٣
تانزانية	٢	٣,١٣
غامبية	٥	٧,٨١
المجموع	٦٤	١٠٠

خامسًا- إجراءات الدراسة:

مـرت الدراسـة الحاليـة بثـلاث مراحـل رئيسـة كـبرى: المرحلـة التحضيريـة، ومرحلـة التطبيـق، ومرحلـة تحليـل البيانـات. ويتخللهـا العديـد مـن المراحـل الصغـرى، وفيمـا يـأتي تفصيـل ذلـك:

١. **المرحلة التحضيرية:**

اسـتغرقت هـذه المرحلـة فصلـين دراسـيين؛ الفصـل الـدراسي الأول، والفصـل الـدراسي الثـاني، مـن العـام الجامعـي ١٤٤٣هـ، وأنجـز فيهـا مـا يـأتي:

١. إجراء مسح للدراسات والبحوث السابقة التي اهتمت بالبحث في تكافؤ الاختبارات المحوسبة مع نظيراتها الورقية، والتي بحثت في الاتجاهات نحو الاختبارات المحوسبة، وإستراتيجيات أداء الاختبارات المحوسبة، للاستفادة من نتائجها في الدراسة الحالية.

٢. تحديد عينة الدراسة، وأخذ الموافقة منهن على المشاركة في الدراسة، عن طريق استمارة الموافقة على المشاركة في الدراسة.

٣. إجراء اختبار التصنيف لعينة الدراسة، وهو اختبار التصنيف المعتمد بجامعة الملك سعود، ثم تقسيمها إلى مجموعتين متكافئتين بناءً على نتائج الاختبار.

٤. تحديد أدوات جمع البيانات، وتصميمها، واتباع عدد من الإجراءات للتحقق من صدقها وثباتها، وهي: (اختبار لغوي -استبانة قياس المعرفة بالحاسب)، وسيأتي تفصيل ذلك في المحور القادم.

٥. أخذ الموافقات وإرسال الخطابات الرسمية للجامعات.

٢. مرحلة التطبيق:

نفذت مرحلة التطبيق في الفصل الدراسي الثاني من العام الجامعي ١٤٤٣هـ، وفيها طبقت أدوات الدراسة على النحو الآتي:

- تطبيق استبانة قياس المعرفة بالحاسب على متعلمات اللغة العربية لغةً ثانيةً (عينة الدراسة).
- تطبيق الاختبار بنسختيه المحوسبة والورقية على عينة الدراسة، إذ قسمت العينة إلى مجموعتين متكافئتين بناءً على اختبار التصنيف، وكل مجموعة قسمت إلى قسمين؛ قسم يؤدي الاختبار المحوسب، وقسم يؤدي الاختبار الورقي، لتحييد جميع الظروف التي قد تؤثر في نتائج الاختبار.

٣. مرحلة تحليل البيانات:

استغرقت هذه المرحلة ثلاثة فصول دراسية، وهي: الفصل الدراسي الثاني من العام الجامعي ١٤٤٣هـ، والفصلان الدراسيان الأول والثاني من العام الجامعي ١٤٤٤هـ. وفيها اتبعت الدراسة عددًا من الإجراءات لتحليل البيانات واستخراج النتائج، ثم تفسيرها ومناقشتها.

أ. تفريغ جميع البيانات وتحليلها.

ب. الوصول إلى النتائج وتفسيرها والإجابة عن أسئلة الدراسة.

- أدوات الدراسة:

وظفت الدراسة أداتين رئيستين لجمع البيانات اللازمة لتحقيق أهدافها، هي: اختبار لغوي، واستبانة قياس المعرفة بالحاسب. وفيما يأتي وصف لهذه الأدوات، وآلية التحقق من صدقها وثباتها:

- **الأداة الأولى- الاختبار اللغوي:**

تهدف هذه الأداة إلى قياس درجة التكافؤ بين الاختبارات المحوسبة والاختبارات الورقية لدى متعلمات اللغة العربية لغةً ثانيةً، وذلك من أجل التحقق من تطابق نتائج الاختبارات المحوسبة مع نتائج نظيراتها الورقية. واستعان الباحثان باختبار اللغة العربية المقنن لغير الناطقين بها (الدليل التدريبي) الصادر عن المركز الوطني للقياس (قياس)، وهو اختبار تحريري يقيس كفايات المتقدمين للاختبار في مهارات الاستماع، والقراءة، والكتابة. وأعد الباحثان نسختين من الاختبار؛ إحداهما ورقية، والأخرى محوسبة على برنامج (Google Forms).

- **استبانة قياس المعرفة بالحاسب:**

صمم الباحثان استبانة لقياس مدى معرفة المتعلمات بالحاسب (ملحق١)، واعتمدا في تصميمها على أدبيات الدراسة، والمراجع ذات الصلة بموضوع الدراسة.

واتبعت الدراسة عددًا من الأساليب للتحقق من صدق الأداة وثباتها، وهي:

- عرض الأداة على عدد من المحكمين المتخصصين في المجال، وإجراء التعديلات اللازمة نحو الملحوظات التي تفضل بها المحكمون.
- معامل ارتباط بيرسون (Pearson) لحساب صدق الاتساق الداخلي لأداة الدراسة.
- حساب الثبات باستخدام معامل الاتساق الداخلي (ألفا كرونباخ).

صدق البناء الداخلي: للتحقق من صدق الاتساق الداخلي لأداة الدراسة الخاصة بقياس مدى المعرفة في الحاسب لمتعلمات اللغة العربية لغةً ثانيةً نحو الاختبارات المحوسبة، حُسِب معامل ارتباط بيرسون بين درجة كل فقرة من فقرات الاستبانة والدرجة الكلية للاستبانة، وجاءت النتائج كما هي موضحة في الجدول الآتي:

جدول رقم (١) معامل ارتباط (بيرسون) بين درجات كل فقرة والدرجة الكلية للاستبانة

م	معامل ارتباط مدى المعرفة في الحاسب
١	**٠,٨١٩
٢	**٠,٧٥٥
٣	**٠,٧٠٣
٤	**٠,٥٦٦
٥	**٠,٣٩٤
٦	**٠,٥٧٠
٧	**٠,٣٩٩
٨	**٠,٦١٥

يتضح لنا من نتائج الجدول السابق أن جميع معاملات ارتباط بيرسون بين درجة كل فقرة من فقرات الاستبانة والدرجة الكلية للاستبانة كانت دالة إحصائيًّا عند مستوى دلالة ($\alpha \leq 0.05$).

ثبـات أداة الدراسـة: اختُـبر ثبـات أداة الدراسـة باسـتخدام معادلـة ألفـا كرونبـاخ (Cronbach's Alpha). ويوضـح الجـدول الآتي معامـلات ثبـات أداة الدراسـة.

جدول رقم (٢) معاملات ثبات أداة الدراسة

المحور	معامل ألفا كرونباخ
ثبات أداة الدراسة	٠,٧٢٧

يتبين من الجدول السابق أن قيم معامل الثبات المقياس هي (٠,٧٢٧).

نتائج السؤال الأول:

مـا مـدى تطابـق الاختبـارات المحوسـبة مـع نظيراتهـا الورقيـة في قيـاس أداء متعلمـات اللغـة العربيـة لغـةً ثانيـةً في الاختبـارات اللغويـة؟

للإجابة عـن هـذا السـؤال قـام الباحثان أولًا بحسـاب المتوسـط الحسـابي والانحراف المعيـاري لمـن خضـن الاختبـار الـورقي والمحوسـب، كما هـو مبين في الجـدول الآتي:

جدول رقم (٣) عدد من خضن الاختبار الورقي والمحوسب والمتوسط الحسابي والانحراف المعياري لدرجات كل مجموعة

طريقة الاختبار	N	Mean	Std. Deviation
ورقي	٣٢	٤٤,٨٤	١٥,٢٧
محوسب	٣٢	٤٧,٣٤	١٣,٨٥

جدول رقم (٤) اختبار «كلوموجروف-سميرونوف» للكشف عن اعتدالية التوزيع للعينتين

One-Sample Kolmogorov-Smirnov Test		
N		64
Normal Parameters	Mean	46.09
	Std. Deviation	14.52
Sig. (2-tailed)		.016

أُجرِي اختبار «كلوموجروف-سميرونوف» للكشف عن اعتدالية التوزيع للعينتين. وأظهرت النتائج أن p-value لاختبار «كلوموجروف-سميرونوف» بلغت (0.016)، ما يعني وجود دلالة إحصائية عند مستوى دلالة (α≤0.05)، أي إن اعتدالية التوزيع للعينتين غير متحققة، ما يلزمنا إجراء الاختبارات اللامعلمية المناسبة في مثل هذه الحالات.

جدول رقم (٥) اختبار «ليفين» للكشف عن تجانس التباين للعينتين

Levene's Test for Equality of Variances			
	F	Sig.	T
Equal variances assumed	0.040	.841	-.686

أُجرِي اختبار «ليفين» للكشف عن تجانس التباين للعينتين. وأظهرت النتائج أن p-value لاختبار «ليفين» بلغت (٠,٨٤١)، ما يعني عدم وجود دلالة إحصائية عند مستوى دلالة (α≤0.05)، أي إن تباين التجانس للعينتين متحقق.

جدول رقم (٦) اختبار «مان-ويتني» لاختبار الفروق بين العينتين المستقلتين

	أداء العينة في الاختبارات
Mann-Whitney U	448.000
Z	-.865
Sig. (2-tailed)	.387
a. Grouping Variable :طريقة الاختبار	

لأن العينتين لا تتبعان التوزيع الطبيعي، فلقد أُجرِي اختبار «مان-ويتني» لاختبار الفروق بين العينتين المستقلتين. وأظهرت النتائج أن p-value لاختبار «مان-ويتني» بلغت (٠,٣٨٧)، ما يعني عدم وجود دلالة إحصائية بين متوسطي العينتين عند مستوى دلالة (α≤0.05).

تفسير نتائج السؤال الأول ومناقشتها:

بالنظر إلى الجدول رقم (٣) يلاحظ أن متوسط درجات المتقدمات للاختبار الورقي من متعلمات اللغة العربية لغةً ثانيةً كانت (٤٤,٨٤)، على حين أن متوسط درجات المتقدمات للاختبار المحوسب كانت (٤٧,٣٤). ويلاحظ من نتائج التحليل أن قيمة الدلالة (0.387=p-value)، أي لا توجد فروق ذات دلالة إحصائية بين درجات متعلمات اللغة العربية لغةً ثانيةً في الاختبارين الورقي والمحوسب عند مستوى دلالة (α≤0.05). وتؤكد هذه النتيجة صدق تكافؤ النتائج في نسختي الاختبار.

وتتفق هذه النتيجة مع دراسة كل من: نيومان وبايدن (Neuman & Baydoun, 1998)، وآل ثاني (٢٠٠٨)، والخزي وزكي (٢٠١١)، وإبراهيمي وزملائه (Ebrahimi, Toroujeni and Shahbazi, 2019)، وياو (Yao, 2020)، التي أظهرت عدم وجود فروق ذات دلالة إحصائية بين درجات المتعلمين في الاختبارات الورقية والمحوسبة. إلا أنها اختلفت عن نتائج دراسة كل من: ديبير وفاسير (Debeer & Visser, 1998)، وحسيني وآخرين (Hosseini et

(Hosseini & Toroujeni, 2017 ،al., 2013)، وحسيني وتوروجيني، التي أظهرت فروقًا لصالح الاختبارات الورقية عند مقارنة نتائجها بالاختبارات المحوسبة. وقد تعارضت أيضًا مع ما توصلت إليه دراسة كلارينا ووالاس (Clariana & Wallace, 2000)، التي أظهرت تفوق الطلبة في الاختبار المحوسب.

ولم تختلف نتائج هذه الدراسة عن نتائج كثير من الدراسات التي طبقت في أمريكا وأوروبا، وعليه فإن نتائج الدراسة تؤكد الثقة العالية أو الثبات العالي الذي تتمتع به الاختبارات المحوسبة، ما يدل على إمكانية الوثوق بنتائج الاختبارات المحوسبة في ميدان تعليم اللغة العربية لغةً ثانيةً. كما تؤكد هذه النتيجة عدم اقتصار استخدام تكنولوجيا الحاسب في المناهج الدراسية على أغراض التدريس فقط، وإمكانية توظيفها في أغراض القياس والتقويم.

وخلاصة ذلك أن تكافؤ نتائج الاختبارات المحوسبة مع نتائج نظيراتها الورقية -الذي توصلت إليه نتائج السؤال الأول-، إضافة إلى المزايا العديدة التي تمتاز بها الاختبارات المحوسبة، ومنها: تقليل تكلفة الطباعة، وتخفيف العبء على المعلم، وتقليل ذاتية التصحيح، يجعل خيار أن تكون الاختبارات المحوسبة بديلًا للاختبارات الورقية خيارًا متاحًا وموثوقًا بنتائجه بناء على دراسة علمية.

نتائج السؤال الثاني:

هل توجد علاقة ارتباطية ذات دلالة إحصائية بين أداء متعلمات اللغة العربية لغةً ثانيةً في الاختبارات المحوسبة وبين خلفيتهن في استخدام الحاسب الآلي؟

للإجابة عن هذا السؤال حُسِب معامل الارتباط «بيرسون» للكشف عن وجود علاقة ارتباطية ذات دلالة إحصائية بين أداء متعلمات اللغة العربية لغةً ثانيةً في الاختبارات المحوسبة وبين خلفيتهن في استخدام الحاسب الآلي، وجاءت النتائج كما هو مبين في الجدول الآتي:

		متوسط الاستجابات
أداء العينة في الاختبار المحوسب	Pearson Correlation	0.066
	Sig. (2-tailed)	0.621

أظهـرت النتـائج أن قيمـة الدلالـة لاختبـار معامـل الارتبـاط «بيرسـون» بلغـت (٠,٦٢١)، مـا يعـني عـدم وجـود دلالـة إحصائيـة عنـد مسـتوى دلالـة (α≤0.05)، أي لا توجـد علاقـة ارتباطيـة ذات دلالـة إحصائيـة بـين أداء متعلمـات اللغـة العربيـة لغـةً ثانيـةً في الاختبـارات المحوسـبة وبـين خلفيتهـن في اسـتخدام الحاسـب الآلي.

السؤال المفتوح:

هل هناك ما تودين إضافته حول مدى معرفتك بالحاسب الآلي؟

كانـت أغلـب إجابـات عينـة الدراسـة تتمحـور حـول إجادتهـن برامـج أخـرى، مثـل: تحريـر مقاطـع الفيديـو، والرسـم علـى برنامـج (Fresh paint)، وحـول دراسـتهن مقـرر الحاسـب الآلي في المسـتوى الرابـع والإضافـة الـتي أحدثهـا في مسـتوى معرفتهـن بالحاسـب الآلي. كما أشـارت إحداهـن إلى أنهـا لا تجيد الكتابة السـريعة على الحاسـب الآلي، وهـذا يفتـح المجـال لدراسـة أخـرى تبحـث في اختبـارات الكتابـة المحوسـبة وعلاقتهـا بسـرعة الكتابـة علـى لوحـة المفاتيـح، عـن طريـق تقنيـات بحثيـة حديثـة، مثـل: تقنيـة تسـجيل النقـرات علـى لوحـة المفاتيـح (keystroke logging)، وبحـث العلاقـة بـين معـدل سـرعة الكتابـة علـى الحاسـب ودرجـات الاختبـار المحوسـب.

تفسير نتائج السؤال الثاني ومناقشتها:

أظهـرت نتـائج السـؤال الثـاني عـدم وجـود علاقـة ارتباطيـة ذات دلالـة إحصائيـة بـين أداء متعلمـات اللغـة العربيـة لغـةً ثانيـةً في الاختبـارات المحوسـبة وبـين خلفيتهـن في اسـتخدام الحاسـب الآلي. ويعـزو الباحثـان هـذه النتيجـة إلى أن متعلمـات اللغـة

العربية لغةً ثانيةً -حسب ما لوحظ في أثناء جمع البيانات- على علم ودراية تامة بكيفية التعامل مع الحاسب الآلي، وخلفيتهن التقنية جيدة على اختلاف جنسياتهن، وقد تعاملن مع الاختبارات المحوسبة خاصة في السنوات الأخيرة، وعليه لم تظهر النتائج ارتباطًا بين المعرفة بالحاسب الآلي والأداء في الاختبار المحوسب، ولعل ما كشفت عنه نتائج السؤال المفتوح من إجادتهن التعامل مع برامج حاسوبية أخرى، مثل: تحرير مقاطع الفيديو والرسم، تعزز هذا التفسير.

وقد تؤكد هذه النتيجة أن المتقدم للاختبار المحوسب لا يحتاج إلى مهارات متقدمة في استخدام الحاسب، وإنما يحتاج إلى الحد الأدنى الذي يمكنه من التعامل مع الحاسب الآلي، فعامل المعرفة بالحاسب الآلي وفق ما خلصت إليه نتائج الدراسة الحالية لا يعد عاملًا مؤثرًا في نتائج الاختبار المحوسب.

وتتفق هذه النتيجة مع نتيجة دراسة كلارينا ووالاس (Clariana & Wallace, 2000) التي خلصت إلى عدم وجود فروق في النتائج تعزى لمتغير المعرفة بالحاسب. كما تتفق كذلك مع دراسة حسيني وآخرين (Hosseini et al., 2013) التي بحثت في العلاقة بين خبرة الطلاب في التعامل مع الحاسب الآلي ودورها في أداء الاختبار المحوسب CBT، وأظهرت نتائجها أن خبرة الطلاب السابقة حول استعمال الحاسب لم يكن لها تأثير كبير في أداء الطلاب في الاختبار المحوسب. وتتفق هذه النتيجة أيضًا مع دراسة حسيني وتوروجيني (Hosseini & Toroujeni, 2017) التي خلصت نتائجها إلى أنه لا يوجد أي ارتباط بين الإلمام بالحاسب ونتيجة الأداء في الاختبار المحوسب.

الخاتمة:

بحثت الدراسة في قضية تكافؤ نتائج الاختبارات المحوسبة مع نظيراتها الورقية في الاختبارات اللغوية، وأثر الخلفية المعرفية بالحاسب الآلي في الأداء في الاختبار المحوسب. ولعل أهم ما خلصت إليه الدراسة، ما يأتي:

- أن الاختبارات المحوسبة متكافئة في نتائجها مع الاختبارات الورقية، ولعل مرجع هذا التكافؤ هو الثبات والصدق العالي الذي تتمتع به الاختبارات المحوسبة، وعليه يمكن عدُّ النسخة المحوسبة من الاختبار بديلًا مناسبًا للاختبار الورقي لمتعلمات اللغة العربية لغةً ثانيةً.
- أن المعرفة بالحاسب الآلي غير مرتبطة بالأداء في الاختبار المحوسب، ولعل ذلك يعود إلى انتشار استخدام الحاسب الآلي انتشارًا واسعًا في مجال التعليم والاختبارات، خاصة بعد جائحة كورونا التي ترتب عليها إقامة اختبارات عن بعد.

توصيات الدراسة:

لعل نتائج هذه الدراسة تقود إلى العديد من القرارات التي تخدم ميدان الاختبارات اللغوية، وعليه يوصي الباحثان باعتماد الاختبارات المحوسبة والوثوق بنتائجها وتفعيلها في المؤسسات التعليمية، نظرًا لثبات تكافؤ نتائجها مع نتائج نظيراتها الورقية، ولما تتمتع به من خصائص، ومنها: مواكبة التطور التكنولوجي. وتوفير نسخ محوسبة من الاختبارات اللغوية إلى جانب النسخ الورقية، وترك حرية الاختيار للمتعلم، بما يعزز مواكبة التطور التقني وحوسبة اختبارات اللغة، ويفتح المجال لمشاريع أكثر تقدمًا في مجال اختبارات اللغة العربية، نحو: الاختبارات التكيفية المحوسبة.

كما توصي الدراسة باستخدام الاختبارات المحوسبة بما يتوافق مع طبيعة المهارات اللغوية، وتدريب معلمي اللغة العربية لغةً ثانيةً على تصميم الاختبارات المحوسبة وتفعيلها والإفادة من مزاياها المتعددة، وتعريض المتعلمات للاختبارات المحوسبة وتدريبهن عليها. والعمل على تطوير أنظمة الاختبارات المحوسبة في المؤسسات التعليمية بما يضمن تلافي المشكلات التقنية.

المقترحات:

في ضوء نتائج هذه الدراسة وتوصياتها يقترح الباحثان إجراء المزيد من البحوث والدراسات التي تعنى بالاختبارات المحوسبة، ومنها:

- إجراء دراسة مماثلة عن تكافؤ الاختبارات المحوسبة ونظيراتها الورقية لدى عينة أخرى، وباستخدام أنواع أخرى من الاختبارات اللغوية، بما يعزز ما أثبتته الدراسة الحالية والدراسات السابقة أو ينفيه من تكافؤ نتائج الاختبارات المحوسبة مع نتائج نظيراتها الورقية في الاختبارات اللغوية.
- إجراء دراسة عن الحاجات التدريبية اللازمة لتفعيل الاختبارات المحوسبة في معاهد ومراكز تعليم اللغة العربية لغةً ثانيةً.

هذه الطبعة إهـداء من المجمـع
ولا يُسمح بنشرها ورقياً أو تداولها تجارياً

المراجع

أولاً-المراجع العربية:

- آل ثاني، العنود. (٢٠٠٨). تكافؤ نموذجي اختبار راف (المحوسب-التقليدي) لدى عينة من طلاب وطالبات جامعة قطر في ضوء قلق الاختبار وقلق الحاسب والاتجاه نحو الحاسب (رسالة دكتوراه غير منشورة). جامعة أم القرى، مكة المكرمة.

- إلويد، دوايت؛ وديفيدسون، بيتر؛ وكوم، كرستين. (٢٠١٤). أساسيات التقييم في التعليم اللغوي، (ترجمة خالد الدامغ). الرياض: مكتبة الملك فهد الوطنية.

- الخزي، فهد، والزكي، محمد. (٢٠١١). تكافؤ الاختبارات الإلكترونية مع الاختبارات الورقية في قياس التحصيل الدراسي: دراسة تجريبية على طلبة كلية التربية بجامعة الكويت. مجلة دراسات الخليج والجزيرة العربية، ٣٧(١٤٣)، ١٩٨-١٦٧.

- الدامغ، خالد. (٢٠١١). الاختبارات والقياس والتقييم معجم موسوعي. الرياض: مدار الوطن للنشر.

- الدامغ، خالد، وعبد الخالق، محمد. (٢٠١٢). الميسر في إعداد الاختبارات لمدرسي اللغات الأجنبية. الرياض: مكتبة الملك فهد الوطنية.

- طه، مهند حسن. (٢٠١٩). فاعلية كل من الاختبارات الإلكترونية والاختبارات الورقية في قياس التحصيل الأكاديمي: دراسة تجريبية على طلاب ماجستير التربية في تكنولوجيا التعليم بجامعة السودان للعلوم والتكنولوجيا. مجلة العلوم التربوية، ٣٠ (١)، ٦٤-٨٢.

- النبهان، موسى. (٢٠٠٤). أساسيات القياس في العلوم السلوكية. عمان: دار الشروق للنشر والتوزيع.

ثانيًا- المراجع الأجنبية:

- AdDamigh, K. (2010). Test Taking Strategies. Riyadh: King Fahd National Library.
- Bennett, R., (2001). How the Internet will help large scale assessment reinvent itself. Education Policy Analysis Archive, 9(5), 1-23.
- Bookman, H. (1989). No more paper and pencil. Personnel Administrator, 26-29.
- Cisar, D., Radosav, D., Markoski, B., Printer, R. & Cisar, P. (2010). Computer adaptive testing of student knowledge. Acta Polytechnic Hungarica, 7(4), 139-152.
- Clariana, R. & Wallace, P. (2002). Paper-based versus computer-based assessment: key factors associated with the test mode effect. British Journal of Educational Technology, 33(5), 593-602.
- Coles, M., Cook, L. & Blake, T. (2007). Assessing obsessive compulsive symptoms and cognitions on the internet: Evidence for the comparability of paper and internet administration. Behavior Research & Therapy, 4(9), 232-240.
- Davidson, P. (2003). Why technology has had only a minimal impact on testing in education. Educational Technology Conference Proceedings, 65-79.
- De Beer, M., & Visser, D. (1998). Comparability of the paper-and-pencil and computerized adaptive versions of the General Scholastic Aptitude Test (GSAT) Senior. South African Journal of Psychology, 28, 21-27.
- Deviant, S. (2011). The Practically Cheating Statistics Handbook. USA: Andale Publishing LLC.

- Ebrahimi, M. R., Toroujeni, S. M. H., & Shahbazi, V. (2019). Score Equivalence, Gender Difference, and Testing Mode Preference in a Comparative Study between Computer-Based Testing and Paper-Based Testing. International Journal of Emerging Technologies in Learning (iJET), 14(07), 128-143.
- Friedrich, S. & Julius, B. (2009). The Transition to Computer-based assessment new approaches to skills assessment and implication for large-scale testing. Luxembourg: Office for Official Publications of the European Communities.
- Hamhuis, E., Glas, C., & Meelissen, M. (2020). Tablet assessment in primary education: Are there performance differences between TIMSS' paper-and-pencil test and tablet test among Dutch grade-four students? British Journal of Educational Technology, 51(6), 2340-2358.
- Hosseini, M., & Toroujeni, S. M. H. (2017). Replacing paper-based testing with an alternative for the assessment of Iranian undergraduate students: Administration mode effect on testing performance. International Journal of Language and Linguistics, 5(3), 78-87.
- Hosseini, M., & Toroujeni, S. M. H. (2017). Validity and Reliability Examination in Onscreen Testing: Interchangeability of Scores in Conventional and Computerized Tests by Examining Two External Moderators. International Journal of English Language Teaching and Linguistics, 1(2), 79-103.
- Hosseini, M., Abidin, M. J. Z., Kamarzarrin, H., & Khaledian, M. (2013). The investigation of Difference between PPT and CBT Results of EFL Learners in Iran: Computer Familiarity and Test Performance in CBT. International Letters of Social and Humanistic Sciences (ILSHS), 11, 66-75.

- Hüseyin, Ö., & Özturan, T. (2018). Computer-based and paper-based testing: Does the test administration mode influence the reliability and validity of achievement tests? Dil ve Dilbilimi Çalışmaları Dergisi, 14(1), 67-85.
- Kearsley, G. (2000). Online education: language and Teaching in Cyberspace. Belmont, CA: Wadsworth.
- Khoshsima, H., Hosseini, M., & Toroujeni, S. M. H. (2017). Cross-Mode Comparability of Computer-Based Testing (CBT) versus Paper-Pencil Based Testing (PPT): An Investigation of Testing Administration Mode among Iranian Intermediate EFL Learners. English Language Teaching, 10(2), 23-32.
- Khoshsima, Morteza., & Toroujeni, S. M. H. (2017). Transitioning to an alternative assessment: Computer-Based Testing and key factors related to testing mode. European Journal of English Language Teaching, 2(1), 54-74.
- Marks, A. & Cronje, J. (2008). Randomized Items in Computer-Based Tests: Russian Roulette in Assessment? Educational Technology & Society, 1(4), 41-50.
- Meijer, R. (2004). Using patterns of summed scores in paper tests and computer adaptive test to detect misfitting item score patterns. Journal of Educational Measurement, 1(2), 199-136.
- Neuman, G. & Baydoun, R. (1998). Computerization of Paper-and- pencil Tests: When ate they equivalent? Applied Psychological Measurement, 22(1), 71-83.
- Russell, M., Goldberg, A., & Oconnor, K. (2003). Computer-based testing and validity: A look back into the future. Assessment in Education: Principles, Policy & Practice, 10(3), 279-293.

- Sim, G. & Horton, M. (2005). Performance and attitude of children in computer based versus paper-based testing. [Online]. Available at: file:///C:/Users/acer/Downloads/paper_20641.pdf
- Woodfield, K. (2003). Getting on board with online testing. THE Journal, 30(6), 32-37.
- Yao, D. (2020). A Comparative Study of Test Takers' Performance on Computer-Based Test and Paper-Based Test across Different CEFR Levels. English Language Teaching, 13(1).

هذه الطبعة إهـداء من المجمع
ولا يُسمح بنشرها ورقياً أو تداولها تجارياً

استعراض لاختبار المستوى الرابع لتخصص اللغة العربية في الجامعات الصينية: واقعه والمأمول منه

د: يه ليانغ ينغ[1]

مستخلص

يعد اختبار المستوى الرابع لتخصص اللغة العربية في الجامعات الصينية الاختبار الموحد الوحيد الذي قام بتنظيمه مجلس التوجيه لتعليم اللغة العربية في الجامعات والمعاهد العالية الصينية التابع لوزارة التربية والتعليم بالصين، والهدف من إقامته هو قياس كفايات الطلبة اللغوية والمهارية وقدرتهم على التطبيق اللغوي المتكامل، الأمر الذي يمنحه مكانة لا يستبدل بها في ضبط العملية التعليمية، وقياس فاعليتها واختبار جودتها. وتقدم هذه الورقة بإيجاز شكل الاختبار وآلية تشغيله، وتُناقش أهميته وتأثيره الاجتماعي، كما تتطرق أيضا إلى كيفية تحسين فاعلية الاختبار، وتعزيز امتثاله لمتطلبات القدرة المنصوص عليها في «المعيار الوطني لتعليم اللغة العربية في الجامعات والمعاهد العالية الصينية»، حتى يؤدي هذا الاختبار الدور المحفز لتدريس اللغة العربية على مستوى الدولة.

الكلمات المفتاحية:

اختبار المستوى الرابع للغة العربية، آلية التشغيل، التأثير الاجتماعي، سبل الإصلاح.

(١) أستاذ بكلية الدراسات العربية بجامعة الدراسات الأجنبية ببكين.

اختبارات اللغة
Journal of Language Testing

Abstract

Test for Arabic Majors Grade Four (TAM-4) in Chinese universities is the only unified exam organized by the Steering Board for Teaching Arabic in Chinese Universities and Higher Institutes of the Ministry of Education in China, and the aim of its establishment is to measure students' linguistic and skill competencies and their ability to apply integrated Arabic language, which gives it an irreplaceable position in controlling the educational process, measuring the teaching effectiveness and testing its quality. This paper briefly presents the form of the test and the mechanism of its operation, discusses its importance and social impact, and also addresses how to improve the effectiveness of the test and enhance its compliance with the language ability requirements stipulated in the "National Standard for Teaching Arabic in Chinese Universities and Higher Institutes", so that this test plays a catalytic role for teaching Arabic at the state level.

Keywords:

TAM-4, Mechanism of employment, Social impact, Ways to improve

هذه الطبعة إهداء من المجمع
ولا يُسمح بنشرها ورقياً أو تداولها تجارياً

المقدمة

مرَّ حتى اليوم أكثر من سبعين عامًا من التطبيق العملي، والبحث العلمي، والتطوير المتواصل، على تدريس اللغة العربية في الجامعات والمعاهد العالية بالصين، منذ أُدْخِلَ رسميًّا في منظومة التعليم العالي الوطني الصيني عام ١٩٤٦. وبعد دخول القرن الحادي والعشرين، ومع إنشاء منتدى التعاون الصيني العربي عام ٢٠٠٤، وتوطيد علاقات التعاون الإستراتيجي الصينية العربية، وخاصة مع طرح مبادرة «الحزام والطريق»[١] ومفهوم «مجتمع ذي مصير مشترك للبشرية»، خطت علاقات التعاون الودية بين الصين والدول العربية عتبة جديدة على أصعدة مختلفة. وتزامناً مع ذلك دخل تعليم اللغة العربية في الجامعات والمعاهد العالية الصينية مسار التطور السريع، وحقق نتائج ملفتة لا مثيل لها. فزاد عدد مؤسسات التعليم العالي المحتضنة لتخصص اللغة العربية إلى ٥٢ جامعة ومعهدًا، مقارنة بسبع جامعات فقط في نهاية القرن العشرين؛ كما تتجاوز مقاعد القبول للطلبة المتخصصين في اللغة العربية وآدابها سنويًّا ألف مقعد، مقارنة مع ١٥٠ مقعدًا فقط في نهاية القرن المنصرم. وفي ظل هذه التطورات المبشرة بالأمل أصبح المستعربون الصينيون يحسون بضرورة إقامة اختبار عام موحد لتقييم فاعلية

(١) مبادرة «الحزام والطريق»، تسمية مختصرة لـ«الحزام الاقتصادي لطريق الحرير» و«طريق الحرير البحري للقرن الحادي والعشرين». ففي سبتمبر وأكتوبر ٢٠١٣، طرح الرئيس الصيني شي جين بينغ أثناء جولته في آسيا الوسطى وجنوب شرقي آسيا، مبادرة البناء المشترك للحزام والطريق مع الدول المعنية. وتغطي هذه المبادرة شرقي آسيا وجنوب شرقي آسيا وجنوبي آسيا وغربي آسيا وآسيا الوسطى ووسط وشرقي أوروبا وغيرها من الدول والمناطق؛ وتتخذ تحقيق «التواصل في السياسات، وترابط المنشآت، وتيسير التجارة، وتيسير التمويل، والتواصل بين قلوب الشعوب» كمحتوى رئيس، و«التشاور المشترك، والبناء المشترك، والاستفادة المشتركة» كمبدأ، و«رابطة المصلحة المشتركة، ورابطة المسؤولية المشتركة، ورابطة المصير المشترك» كأهداف، سعيًا إلى إسعاد الدول والشعوب التي تشارك في البناء المشترك لها.

التدريس، وقياس مستوى الطلبة في تطبيق اللغة العربية، والتوصل إلى نقاط القوة في عملية التدريس للتمسك بها، وإلى نقاط الضعف لتداركها مستقبلا، حتى تشهد قضية تعليم اللغة العربية في الصين تطورًا سلسًا منضبطًا ومتوازنًا.

١ - نشأة اختبار المستوى الرابع لتخصص اللغة العربية بالجامعات الصينية

سعيًا إلى ضبط عملية التدريس للغة العربية في مؤسسات التعليم العالي الصينية، وجعلها تؤسس على قاعدة علمية متقدمة، والوصول بها إلى مستوى تعليمي مرض، قام المستعربون الصينيون بصياغة «منهج تعليم اللغة العربية في الجامعات الصينية» (يشار إليه فيما يلي باختصار «المنهج»)، وإصداره رسميًّا عام ٢٠٠٠. يعد المنهج تلخيصًا تاريخيًّا شاملًا لتعليم اللغة العربية في الصين عبر السنين، وهو منهج متكامل يستهدف تعزيز مستوى التعليم والبحوث الأكاديمية للغة العربية في الجامعات الصينية، وهو أيضا وثيقة توجيهية تنظم بشكل علمي نشاطات التعليم في مختلف الجامعات، حيث أوضح لكل فصل من التدريس ما يخصه من المؤشرات كمًا ونوعًا لكلا الجانبين المعرفي والكفائي، بل حدد لكل مرحلة دراسية غاية التدريس ومبادءه ومحتوياته وترتيباته والتزاماته، وكيفية تقييم مستوى الطلبة فيها. وهو بذلك يتحلى بتوجيه مبدئي عام مع مؤشرات كمية تفصيلية، وفي الوقت نفسه يجمع هذا المنهج بين وحدة المطالب ومرونة الترتيب، ويهتم بتدريب المهارات اللغوية الأساسية، وأيضا بتكوين القدرات الشاملة عند الطلاب. وفيما يخص التقييم، حدد المنهج على نحو تفصيلي عناصر قياس مختلف المهارات والكفاءات ومحتوياتها وأساليبها ومتطلباتها وأغراضها، ووقت الاختبار لكل عنصر ونسبته من الكل، (المنهج، ٢٠٠٠) مما يوفر معيارًا قياسيًّا موحدًا لتقييم فاعلية طرق التدريس عند المؤسسات التعليمية المختلفة. وبفضل تضافر الجهود من قبل المستعربين الصينيين، تم في العام نفسه تأليف «جدول

هذه الطبعة إهداء من المجمع
ولا يُسمح بنشرها ورقياً أو تداولها تجارياً

المفردات لمنهج تعليم اللغة العربية في الجامعات الصينية» وإصداره، الذي يضم ٩ آلاف مفردة و٣ آلاف عبارة اصطلاحية. وقُسِّمتْ هذه المفردات والعبارات إلى ثمانِ درجات في كتاب الجدول حسب نسبة تداولها ومدى صعوبتها في الاستيعاب، الأمر الذي جعل إجراء اختبار وطني علميّ موحد أمرًا ممكنًا.

في عام ٢٠٠٦، ومن أجل مراجعة مدى تنفيذ تخصص اللغة العربية في المؤسسات التعليمية الجامعية للمنهج، وجعله يطبّق على أرض الواقع بكل معنى الكلمة، وكذلك من أجل تقييم أداء الطلاب في العملية الدراسية على نحو موضوعي وشامل وعادل، وتقييم جودة التدريس في الجامعات على اختلافها من حيث تاريخ التأسيس، وحجم هيئة التدريس، وتراكم الخبرات التعليمية، تم إنشاء فريق خبراء اختبار المستوى الرابع لتخصص اللغة العربية في الجامعات الصينية (يشار إليه فيما يلي باختصار «فريق الخبراء»)، وذلك بالتحفيز والقيادة المباشرة من قبل مجلس التوجيه لتعليم اللغة العربية في الجامعات والمعاهد العالية الصينية التابع لوزارة التربية والتعليم بالصين (يشار إليه فيما يلي باختصار «مجلس التوجيه»).

قام فريق الخبراء بدراسة معمّقة لآخر التطورات في النظريات اللغوية والتعليمية ونظريات الاختبارات، وباستقصاء منظم لتجارب مختلف المؤسسات التعليمية في تنظيم الاختبارات، وأجرى استطلاعات طالت نواحي عديدة، وجمع معطيات مباشرة موثوقًا بها؛ وفي الوقت عينه حرص فريق الخبراء على الاستفادة من آراء الخبراء المحنكين محليًّا أو أجنبيًّا، ووضع على أساس ذلك التصميم الإطاري للاختبار، بحيث يشمل هدف الاختبار وخاصيته ونطاقه، وطريقة تنظيمه وإجرائه، وشكله، ومحتوياته. وتحت جهود هادفة، وإتقان محكم، بدأ فريق الخبراء بصياغة «مسودة منهاج اختبار المستوى الرابع لتخصص اللغة العربية في الجامعات والمعاهد العالية الصينية». وبعد ذلك نظم فريق الخبراء في كل من عامي ٢٠٠٧ و٢٠٠٨ اختبارين تجريبيين على مستوى

الدولة، شارك فيهما جميع الطلبة المتخصصين في اللغة العربية الذين أكملوا المرحلة الدراسية الأساسية مستعدين للانتقال إلى المرحلة الدراسية المتقدمة في مختلف الجامعات والمعاهد العالية. وقد بلغ عدد المتقدمين إلى الاختبار في عام ٢٠٠٧، أربعمائة طالب ونيف من ١٣ جامعة ومعهدًا. وبعد التحليل الإحصائي العلمي لنتائج الطلبة في الاختبار، والاستماع إلى ردود الأفعال من قبل المدرسين والممتحَنين، قام فريق الخبراء بإدخال تعديلات على مسودة منهاج الاختبار، فأضاف اختبار الاستماع إليه؛ ليكون مجموع الدرجات ١٢٠ درجة. وبعد اختبار عام ٢٠٠٨، وعلى أساس التحليل الاستقصائي لنتائج الطلبة في مختلف العناصر، وتجميع الآراء والمقترحات، قام فريق الخبراء بضبط أنواع الأسئلة، ومقدار كل نوع منها، وإقرار منهاج الاختبار؛ ليخرج بشكله النهائي العلمي اللامع.

خلاصة القول، إن اختبار المستوى الرابع لتخصص اللغة العربية بالجامعات الصينية هو الاختبار الوحيد الذي قام بتنظيمه «مجلس التوجيه»، وهو الاختبار العام الذي يتقدم إليه كل الطلبة الصينيين المتخصصين في اللغة العربية بعد أن يكملوا دراستهم للغة العربية سنتين، وهو الاختبار المعياري الذي يصمم لقياس ما إذا كان الطلبة يقدرون على الارتقاء إلى المستوى الدراسي المتقدم بعد أن يكملوا المرحلة الدراسية الأساسية.

٢ - وضع اختبار المستوى الرابع لتخصص اللغة العربية في الجامعات الصينية

مرت ١٣ سنة على إقامة اختبار المستوى الرابع لتخصص اللغة العربية في الجامعات الصينية (يشار إليه فيما يأتي باختصار «الاختبار العام») منذ إطلاقه عام ٢٠٠٩. ويُعَدّ اختبارًا معياريًّا وطنيًّا موحدًا ووحيدًا لتحديد مستوى الطلبة الصينيين المتخصصين في اللغة العربية بعد دراستهم للعربية لمدة سنتين، كان الهدف من إقامته هو مراجعة مدى تنفيذ تخصص اللغة العربية في المؤسسات

هذه الطبعة إهداء من المجمع
ولا يُسمح بنشرها ورقياً أو تداولها تجارياً

التعليمية الجامعية للمنهج، والكشف عن نقاط الضعف في عملية التعليم؛ حتى يتاح للمؤسسات التعليمية تبني إجراءات علمية لازمة لتداركها. ويشمل نطاق الاختبار جميع المعارف اللغوية التي حُدِّدتْ في منهج المرحلة الأساسية، والمهارات اللغوية الأربع، وهي فهم المسموع، وفهم المقروء، والتعبير الكتابي، والترجمة التحريرية بين اللغتين الصينية والعربية.

ولضمان تحقيق الطبيعة العلمية والموضوعية والموثوقية للاختبار العام، مع مراعاة جدوى وخصائص الاختبار الذي يتقدم إليه سنويًّا ما يزيد عن ألف طالب وطالبة، وكذلك لأجل ضمان مصداقيته، تأتي معظم الأسئلة الواردة في الاختبار على شكل أسئلة موضوعية أي الاختيار من متعدد، باستثناء الجزء الخاص بالترجمة التحريرية والإنشاء، الذي تأتي الأسئلة فيه مقالية لتتيح للطالب الحرية، أو الحرية المشروطة في كتابة الإجابة التعبيرية والإنشائية، والقيام بصياغتها بالطريقة التي يريدها، مما يعطي الطالب مجالًا للتعبير عن أفكاره مستعينًا بقدرته على إنشاء جمل مترابطة، ومقال متسلسل. وفيما يلي جدول يبين أقسام الاختبار العام، وعدد الأسئلة في كل قسم، والدرجات والوقت المحدد لأدائها.

الجدول الأول: تكوين الاختبار

القسم	الأول	الثاني	الثالث	الرابع	الخامس	المجموع
المحتوى	الاستماع	الصرف والنحو	فهم المقروء	الترجمة	الإنشاء	
عدد الأسئلة	٢١	٨٠	٣٦	١٦	٢	١٥٥
الدرجات	٢٠	٤٠	٢٥	٢٠	١٥	١٢٠
وقت الأداء	٣٠ دقيقة	٩٠ دقيقة		٦٠ دقيقة		١٨٠ دقيقة

يوضح لنا الجدول أن الاختبار العام يستمر ثلاث ساعات، ومجموع الدرجات هو ١٢٠ درجة، ويبلغ عدد الأسئلة فيه ١٥٥ سؤالًا. ويتكون الاختبار من خمسة أقسام تضم الاستماع، والمفردات، والصرف، والنحو، وفهم المقروء، والترجمة التحريرية، والإنشاء. وأسئلة الأقسام الثلاثة الأولى موضوعية، وتأخذ ٨٥ درجة، وتمثل ٧٠،٨٪ من الدرجات الكلية؛ أما القسمان الأخيران فأسئلتهما مقالية، ويأخذان ٣٥ درجة، ويمثلان ٢٩،٢٪ من الدرجات الكلية. ومن حيث محتوى الأسئلة، ينقسم الاستماع إلى جزأين، هما الإملاء وفهم المسموع، ويأخذان ٢٠ درجة، ويمثلان ١٦,٧٪ من مجموع الدرجات؛ وتنقسم الأسئلة الهادفة لقياس حصيلة الطالب المعرفية باللغة العربية إلى ثلاثة عناصر، وهي المفردات (أسئلتها اختيار المترادفات والأضداد للكلمات أو العبارات حسب معانيها في سياق الكلام) والصرف والنحو، وتبلغ درجاتها ٤٠ درجة حصيلة الإجابة عن ٨٠ سؤالًا، وتمثل ٣٣،٣٪ من مجموع الدرجات؛ وتنقسم الأسئلة الهادفة لقياس قدرة الطالب على التطبيق الشامل للعربية أيضا إلى ثلاثة عناصر، وهي فهم المقروء، والترجمة التحريرية، والإنشاء، وتبلغ درجاتها ٦٠ درجة، وتمثل ٥٠٪ من إجمالي الدرجات، مما يؤكد لنا أن الاختبار العام يولي عناية خاصة بقدرة الطالب على استخدام اللغة العربية حسب المناسبات.

تُحسب التقديرات التي يحصل عليها الطلبة على النحو التالي: تقدير «مقبول» من ٥٥ حتى ٧٩،٩ درجة؛ وتقدير «جيد جدا» من ٨٠ حتى ٩٤،٩ درجة؛ وتقدير «امتياز» وهو ما فوق ٩٥ درجة؛ و»الفشل» دون ٥٥ درجة. فيما يلي إحصائيات نتائج الطلبة في الفترة ما بين عامي ٢٠٠٩ و٢٠٢١.

هذه الطبعة إهداء من المجمع
ولا يُسمح بنشرها ورقياً أو تداولها تجارياً

الجدول الثاني: إحصائيات نتائج الطلبة في الفترة ما بين عامي ٢٠٠٩ و٢٠٢١

العام الدراسي	عدد المشاركين	نسبة "مقبول"	نسبة "جيد جدا"	نسبة "امتياز"
٢٠٠٩	٤٩٠	٢٥،٤٪	٣٧،٣٪	١٢،٢٪
٢٠١٠	٥٨٢	٢٩،٢٪	٢٤،٩٪	١٥،٨٪
٢٠١١	٦٤٩	٣٤،٣٪	١٧،٩٪	٥،١٪
٢٠١٢	٧٦٥	١٢،١٪	١٩،٩٪	٩،٣٪
٢٠١٣	٩١٠	٣٩،٦٪	١٦،٥٪	٦،٥٪
٢٠١٤	١٠٤٣	٣٦،٣٪	١٧،٣٪	٥،٠٪
٢٠١٥	١٠٨٧	٣٤،٨٪	١١،٥٪	٥،١٪
٢٠١٦	١١٩٩	٢٩،٢٪	١٣،٨٪	٤،٨٪
٢٠١٧	١٢٥١	٣٣،٩٪	٩،٨٪	٢،٦٪
٢٠١٨	١٣٣٥	٣٥،٦٪	٩،٣٪	٣،٦٪
٢٠١٩	١٢٢٣	٣٨،٨٪	١٨،٦٪	٨،٨٪
٢٠٢١	١٢٣٥	٢٤،٠٪	٩،٤٪	٣،٣٪

ومن البيانات الواردة في الجدول أعلاه، يمكن ملاحظة أن عدد المشاركين في الاختبار أظهر اتجاهًا تصاعديًّا في السنوات العشر الأولى؛ ليصل إلى ذروته عام ٢٠١٨ برقم ١٣٣٥، وبعد ذلك بدأ يشهد تراجعًا طفيفًا، ويستقر عند ألف ومائتين ونيف. وبالمقارنة بين عدد المشاركين بين عامي ٢٠٠٩ و٢٠٢١، نلاحظ أنه تضاعف ٢،٥٢ مرة، مما يثبت أن تعليم اللغة العربية في الصين شهد ازدهارًا لا مثيل له في العقد الثاني من القرن الجديد. وبسبب تأثير جائحة كورونا، لم يتم تنظيم الاختبار عام ٢٠٢٠. ومما يبيّنه الجدول أن نسبة الطلبة الناجحين في الاختبار تراجعت بشكل ملحوظ بين عامي ٢٠٠٩ و٢٠٢١، لتهبط من ٧٤،٩٪ (أي ١٢،٢٪ + ٣٧،٣٪ + ٢٥،٤٪) إلى ٣٦،٧٪ (أي ٣،٣٪ + ٩،٤٪ + ٢٤،٠٪)، وكذلك نسبة الطلبة

الحاصلين على تقدير «جيد جدا» تراجعت من ٣٧،٣٪ عام ٢٠٠٩ إلى ٩،٤٪ عام ٢٠٢١، أما نسبة الطلبة الحاصلين على تقدير «امتياز»، فقد تقلصت من ١٢،٢٪ عام ٢٠٠٩ إلى ٣،٣٪ عام ٢٠٢١. والأسوأ من ذلك أن نسبة الذين حصلوا على نتائج جيدة، أي ما فوق ٨٠ درجة من الإجمالي البالغ ١٢٠ درجة، تراجعت من ٤٩،٥٪ (أي ٣٧،٣٪ + ١٢،٢٪) عام ٢٠٠٩ إلى ١٢،٧٪ (أي ٩،٤٪ + ٣،٣٪) عام ٢٠٢١. وإن دلت هذه البيانات على شيء، فإنما تدل على أنه مع الزيادة في حجم القبول للطلبة المتخصصين في اللغة العربية، لم تشهد نوعية التعليم ما يلزم من الاستقرار، ناهيك عن الارتقاء، مما يدق ناقوس الخطر من انخفاض مستوى تعليم اللغة العربية بالصين.

٣ - الاختبار العام للغة العربية ومستوى تعليم اللغة العربية في الصين

ترى نظرية التدريس أن دورة التدريس الكاملة تتضمن الحلقات الخمس التالية: تحليل مهام التدريس، وتصميم منهج التعليم، وتأليف الكتب المدرسية، وتنظيم التدريس وتنفيذه، والاختبار. (شيو، ٢٠٠٤) مما يؤكد أن الاختبار حلقة لا غنى عنها في سلسلة التدريس، إذ إنه لا يقيس فقط استيعاب الطلبة للمعارف والمهارات والتطبيق العملي للغة التي يدرسونها فحسب، بل يوفر أيضا البيانات الموثوق بها للمعلمين، أو المؤسسات التعليمية؛ للتعرف على حالة التدريس وفاعليته.

والاختبار العام لا يتقيد بمواد التدريس، وإنما يُعَد وفق ما تم تحديده من المعايير في منهاج الاختبار وتشرف على إعداد أسئلته فرقة الخبراء. ويتم تقييم الطلبة، وتصنيفهم حسب نتائجهم فيه، فالذين نالوا أكثر من ٥٥ درجة، يصنّفون من الناجحين؛ فيحصل على شهادة بتقدير «مقبول» الذين نالوا درجات من ٥٥ حتى ٧٩،٩ درجة؛ ويحصل على شهادة بتقدير «جيد جدا» الذين نالوا من ٨٠ حتى

هذه الطبعة إهداء من المجمع
ولا يُسمح بنشرها ورقياً أو تداولها تجارياً

٩٤،٩ درجة؛ ويحصل على شهادة بتقدير «امتياز» الذين نالوا ما فوق ٩٥ درجة. والمكتب الوطني الصيني لتنظيم اختبارات اللغات الأجنبية مسؤول عن إعداد وتوزيع الشهادات. أما الذين لم يجتازوا الاختبار في السنة الدراسية المحددة، فلا تتاح لهم إلا فرصة واحدة فقط للمشاركة في الاختبار التكميلي، وإذا وفّقوا فيه، حصلوا فقط على شهادة بتقدير «مقبول». لذا يولي الطلبة جميعهم في تخصص اللغة العربية في المؤسسات التعليمية الصينية اهتمامًا كبيرًا لهذا الاختبار، كما يمكن لنتائج الاختبار أن تقيّم بشكل موضوعي الوضع التعليمي الصيني الحالي للغة العربية. فمن خلال تحليل النتائج، يمكننا الإلمام بما تم تحقيقه من النتائج، كما يمكننا أيضًا الإحاطة بالنقص والمشاكل التي تواجه عملية التدريس. وبالتالي يتسنى لصانعي القرار اتخاذ إجراءات يرونها ضرورية لتحسين جودة التعليم، من خلال تعديل متطلبات المقررات، ومراجعة خطة التدريس، وتأليف الكتب المدرسية الملائمة، وتحسين طرق التدريس وأساليبه. فقد أشار لايل ف. باشمان (Lyle F. Bachman)، الخبير في الاختبارات اللغوية، والأستاذ بجامعة كاليفورنيا (لوس أنجلوس)، إلى أن «صحة القرار وسلامته في المجال التعليمي تتوقفان من ناحية على قدرة صانعي القرار وكفاءتهم، وعلى البيانات التي تم جمعها من ناحية أخرى. ففي ظل تساوي كافة الظروف الأخرى، كلما كانت المعلومات التي تم جمعها ذات موثوقية أكبر وذات صلة أوثق، كلما زادت إمكانية الوصول إلى القرار الصحيح والسليم.» (ليو، هان، ٢٠٠١) واستنادًا إلى ذلك، فإن الخطوة الأولى لتقييم وضع التعليم الصيني للغة العربية بشكل موضوعي، هي الحصول على معلومات موثوق بها. ويوفر الاختبار العام بمصداقيته وموثوقيته وشموليته إمكانية الوصول إلى الهدف. وفيما يلي نقوم بمراجعة شاملة لنتائج الطلبة في الاختبارات من عام ٢٠٠٩ حتى عام ٢٠٢١ حسب العناصر المتحَنة، حتى نتوصل إلى بعض الاستنتاجات بشأن وضع تعليم اللغة العربية في الصين.

اختبارات اللغة
Journal of Language Testing

الجدول الثالث: متوسط درجات الطلبة في الاختبار حسب العناصر الممتحَنة

السنة الدراسية	الإملاء (١٠ درجات)	فهم المسموع (١٠ درجات)	المفردات والصرف والنحو (٤٠ درجة)	كلوز (٥ درجات)	فهم المقروء (٢٠ درجة)	الترجمة بين اللغتين (٢٠ درجة)	الإنشاء (١٥ درجة)	المجموع (١٢٠ درجة)
٢٠٠٩	٣،٠٧	٦،٩٧	٢٣،٩٤	٢،٧٦	١٣،٢٣	١٠،٢٧	٨،٠١	٦٨،٢٥
٢٠١٠	٤،٥٣	٦،٣٥	٢٤،٣٣	٢،٥٨	١٤،١٩	٩،٩١	٧،٣٩	٦٩،٢٨
٢٠١١	٥،١١	٥،٥٦	٢١،٦٦	٢،١٨	٩،٧١	٩،٠٨	٧،٢١	٦٠،٥١
٢٠١٢	٢،٩١	٦،٧٥	٢٢،٩٠	٢،١٣	١٢،١٢	٨،٩٢	٦،٦٠	٦٢،٣٣
٢٠١٣	٤،٤١	٧،٢٤	٢١،٥٧	١،٧٨	١١،١٩	٨،٦٠	٧،٢٦	٦٢،٠٥
٢٠١٤	٤،٦٠	٥،٨٨	٢١،٨٣	١،٤٧	١١،١٧	٨،٨٩	٦،٦٠	٦٠،٤٤
٢٠١٥	٤،١١	٥،٠٠	٢١،٩٩	٢،٠٦	١٠،١٠	٧،٤٣	٦،٢٤	٥٦،٩٣
٢٠١٦	٣،٦٨	٥،٠٧	١٩،٨٨	٣،١٣	١٠،٦١	٧،٦٩	٥،٢٩	٥٥،٣٥
٢٠١٧	٣،٤٦	٥،٤٩	١٩،٦١	١،٦٦	٩،٦٣	٧،٥٠	٥،٩٦	٥٣،٣١
٢٠١٨	٣،٠٢	٥،٦٤	١٩،٢٦	٢،٢٠	٩،٢٩	٨،٢٢	٦،٠٥	٥٣،٦٨
٢٠١٩	٤،٤٧	٥،٨٣	٢١،٤٤	٢،٢٨	١١،١٨	١١،٨٧	٧،٥٤	٦٤،٦١
٢٠٢١	٣،٠٥	٥،١٥	١٨،٢٥	١،٦١	٩،٢٨	٨،٨٨	٦،٥٩	٥٢،٨١
متوسط الدرجات	٣،٨٧	٥،٩١	٢١،٣٩	٢،١٥	١٠،٩٨	٨،٩٤	٦،٧٣	٥٩،٩٦

هذه الطبعة إهداء من المجمع
ولا يُسمح بنشرها ورقياً أو تداولها تجارياً

وتشير البيانات الموضحة في جدول متوسط درجات الطلبة للعناصر السبعة التي تم قياسها إلى ما يلي:

١. العنصر الذي حقق فيه الطلبة أدنى معدل هو الإملاء. ففي عام ٢٠١٢ الذي سجل فيه الطلبة أدنى الدرجات، كان معدل الدرجات هو ٢٩،١٪ فقط؛ وحتى في عام ٢٠١١ الذي سجل فيه الطلبة أعلى الدرجات، كان معدل الدرجات هو ٥١،١٪. فلا تبشّر هذه النتائج بالتفاؤل لحالة تعليم اللغة العربية في الصين، إذ ترى نظرية الاختبار الحديثة أن «الإملاء يُعَد عنصرًا ذا شمولية في الاختبار، لا يختبر فقط المعارف اللغوية عند الطلبة، بل يختبر أيضًا قدرتهم على التنبؤ من خلال الاستعانة بالسياق اللغوي، والمعارف المتراكمة والحس اللغوي عندهم، وذلك إلى جانب أنه يختبر على نحو شامل مستوى إتقانهم للغة.» (وو، ٢٠٠٤) و«الإملاء هو اختبار شامل يقيس قدرة الطلبة على استخدام اللغة، حيث إنه لا يقيس ما عند الطلبة من العناصر اللغوية والمهارات اللغوية، مثل كمية المفردات، والقدرة على بناء الجملة، والقدرة على فهم المسموع والمقروء فحسب، بل يقيس أيضًا قدرتهم على إنشاء المقالة. وهذا يعني أن الذين يحسنون في الإملاء يملكون كفاءة تواصل كافية في عملية التواصل اللغوي الطبيعي.» (ليو، هان، ٢٠٠٤) واستنادًا إلى نظرية الاختبار الحديثة هذه، لا تزال قدرة الطلبة على التواصل الفعال باللغة العربية في حاجة ملحة للتعزيز.

٢. العنصر الذي حصل فيه الطلبة على ثاني أدنى درجة هو كلوز (cloze). فمتوسط درجات الطلبة في الأعوام الاثني عشر التي أُجرِي فيها الاختبار هو ٢،١٥ درجة من مجموعه خمس درجات، مما يعني أن المعدل المئوي لنتيجة الطلبة هو ٤٣٪. وفي عام ٢٠١٤ الذي سجل فيه الطلبة أدنى

الدرجات، كان المعدل المئوي لنتيجتهم هو ٢٩،٤٪ فقط. والمعروف أن اختبار كلوز يقيس قدرة الممتحَنين على فهم السياق والمفردات من أجل تحديد الكلمات الصحيحة، أو نوع الكلمات التي تناسب المكان الفارغ في النص؛ وهو لا يقيس مدى استيعاب الممتحَنين لاتّساق معنى النص وترابطه، والاستخدام السليم والملائم للغة فقط، بل يقيّم أيضا مدى قدرتهم على الفهم الدقيق للسياق، وصدق الحكم المنطقي لديهم. لذا فإنه نوع من الأسئلة التي تتصف بالشمولية والمرونة القوية، تتطلب صحة الإجابة عنه رفعة في الذوق اللغوي. وفي سبعينيات القرن الماضي، طرح اللغويون الأمريكان على رأسهم السيد جون دبليو أولر (John W. Oller) «فرضية الكفاءة اللغوية الشاملة، واكتشفوا بعد الكثير من الدراسات أن بعض أساليب الاختبار الشامل التي يمثلها الإملاء واختبار كلوز يمكن أن تقيس بشكل أفضل المستوى اللغوي الحقيقي عند الطلاب.» (ليو، هان، ٢٠٠٤) ويبين لنا الجدول السابق أن الإملاء واختبار كلوز أصبحا أبرز نقطتي ضعف عند الطلاب الممتحَنين الصينيين، والأمر الذي يظهر أن المستوى اللغوي الحقيقي عند الطلاب محدود، وفي حاجة إلى التعزيز.

٣. العنصر الذي حصل فيه الطلبة على ثالث أدنى درجة، هو الترجمة التحريرية بين اللغتين الصينية والعربية. فمتوسط درجات الطلبة في الأعوام الاثني عشر التي أُجري فيها الاختبار هو ٨،٩٤ درجة من مجموع ٢٠ درجة، مما يشير إلى أن المعدل المئوي لنتيجة الطلبة هو ٤٤،٧٪. وفي عام ٢٠١٥ الذي سجل فيه الطلبة أدنى الدرجات، كان متوسط الدرجات هو ٧،٤٣ درجة، ما يعني أن المعدل المئوي لنتيجتهم هو ٣٧,١٥٪ فقط. ومن المسلم به أن الترجمة عملية إبداعية بين لغتين، يتطلب تحسين الكفاءة فيها تنمية الرصيد اللغوي والمعرفي، وحسن توظيف المفردات في سياقات مناسبة، ولا يتحقق ذلك إلا بعد تراكم لغوي متواصل، وتدريب واع مستمر.

٤. يأتي بعد ذلك الإنشاء حسب متوسط درجات المتحَنين بالترتيب التصاعدي. فمتوسط درجات الطلبة في الأعوام الاثني عشر هو ٦،٧٣ درجة من مجموع ١٥ درجة، مما يشير إلى أن المعدل المئوي لنتيجتهم هو ٤٤،٩٪. وفي عام ٢٠١٦ الذي سجل فيه الطلبة أدنى الدرجات، كان متوسط الدرجات هو ٥،٢٩ درجة، ما يعني أن المعدل المئوي لنتيجتهم في هذا العنصر هو ٣٥،٣٪ فقط. ويتكون الإنشاء في الاختبار العام من قسمين: القسم الأول يأخذ خمس درجات، وهو عبارة عن كتابة مقال تطبيقي بسيط حسب المطلوب، على ألا يقل عدد الكلمات فيه عن ثلاثين كلمة عربية، ومن أنواعه: الدعوة، والاعتذار، والتهنئة، والإعلان عن المحاضرة أو الاجتماع. والجزء الثاني يأخذ عشر درجات، وهو عبارة عن اختيار في موضوع من موضوعين لكتابة مقالة لا تقل كلماتها عن مئة كلمة عربية.

يُعَد الإنشاء أحد أساسيات دراسة اللغة، ويؤدي دورًا لا غنى عنه في زيادة مهارات الطالب في الإفصاح عن نفسه، وطرح أفكاره، وصياغة آرائه ومقترحاته، وتبلور وجهات نظره، مستظهرًا من خلال ذلك قوة صياغته، وفصاحة لسانه، وجمال ألفاظه. ويقيس الإنشاء في هذا الاختبار قدرة الطلبة على التعبير عما يرونه، وما يفكرون فيه، وما يعيرون له اهتمامهم في حياتهم بلغة عربية سليمة متماسكة، وتفكير منظم، وأسلوب واضح. ومن خلال المقارنة بين درجات الطلبة في هذا العنصر يمكننا ملاحظة الفجوة الكبيرة بين مختلف المتحَنين في مدى استيعابهم للمعارف اللغوية، وخصائص الأنماط النصية، والتطبيق العملي للغة.

٥. إلى جانب العناصر المذكورة أعلاه، غالبًا ما يكون متوسط درجات المتحَنين بالترتيب التصاعدي كما يلي: يأتي أولًا فهم المقروء، وعلى إثره فهم المسموع، وأخيرًا المفردات والصرف والنحو. وهناك تذبذبات في نتائج المتحَنين لعنصري «فهم المقروء» و«فهم المسموع» من عام إلى آخر،

بينما يستقر أداؤهم في عنصر «المفردات والصرف والنحو»، وسجلوا أعلى الدرجات في كل عام من الأعوام الاثني عشر، مما يشير إلى أن تخصص اللغة العربية في الجامعات والمعاهد العالية الصينية يلجأ عادةً إلى "تعليمية الألفاظ والقواعد النحوية" بوصفها وسيلةً من وسائل تعليم اللغة العربية.

تظهر البيانات والتحليلات أعلاه أنه بعد توسيع نطاق شروط القبول للطلبة في تخصص اللغة العربية، يعاني عدد غير قليل منهم من نقص في جانب أو أكثر من جانب في تعلم اللغات الأجنبية، وهي كفايات الاستماع، والتعبير الشفوي، والقراءة، والتعبير الكتابي، والترجمة، وهم في حاجة ملحة لتعزيز القدرة على تطبيق اللغة شفويًا وكتابيًا، والاستخدام الدقيق والملائم للألفاظ لمقتضى الحال. وهذا يعني أن تخصص اللغة العربية في الصين ما زال أمامه طريق طويل شاق ليقطعه قبل تحقيق جودة التعليم المرضية.

٤ - التأثير الاجتماعي للاختبار العام

يعدّ اختبار المستوى الرابع لتخصص اللغة العربية في الجامعات الصينية اختبارًا معياريًا، فبدلًا من المقارنة بين حصيلة طالب وآخر، يهدف إلى معرفة ما إذا كان الطلبة قد اكتسبوا المهارات المعرفية واللغوية اللازمة للارتقاء إلى الفترة الدراسية الجامعية المتقدمة، بعد أن يكملوا دراستهم لتخصص اللغة العربية في الفترة التمهيدية. ويعد الاختبار العام هو الاختبار المعياري الوحيد الذي يقيمه تخصص اللغة العربية تحت إشراف مجلس التوجيه التابع لوزارة التربية والتعليم الصينية، ويشارك فيه جميع الجامعات والمعاهد العالية التي تحتضن تخصص اللغة العربية؛ وتعده جميع الجامعات اختبارًا عامًا ذا فاعلية أكبر في قياس جودة التعليم، وفرصة سانحة للتعرف على مدى امتلاك الطلبة للمهارات والمعارف اللغوية الأساسية، وتحديد جوانب القوة والضعف لديهم،

هذه الطبعة إهداء من المجمع
ولا يُسمح بنشرها ورقياً أو تداولها تجارياً

وكذلك توفير بيانات موثوق بها تساهم في دعم اتخاذ القرار على مستوى الكليات أو المؤسسات التعليمية. وذلك لأن هذا الاختبار «له مجموعة متكاملة ومترابطة من الأنظمة في صياغة أسئلة الاختبار وتنفيذه، وتسجيل الدرجات وتفسيرها. وبهذا فقط، يتميز الاختبار بمعيار موحّد، وبالتالي تبرز قابليته للمقارنة بين نتائج مختلف المتحنين.» (ليو، وهان، ٢٠٠٤)

والهدف من إعداد الاختبار هو مراجعة تنفيذ الجامعات المختلفة للمنهج في مرحلة التعليم الأساسية للغة العربية، وتقييم فاعلية العملية التعليمية، وتعزيز التواصل بين الجامعات، ودفع الإصلاح، وتحسين الجودة من خلال مواصلة التنفيذ الفعال للمنهج. فقد شهد الاختبار العام نضجًا وتطورًا ملحوظين في كل من المحتوى والشكل، بعدما تراكمت الخبرات من خلال إقامة الاختبار التجريبي في عامي ٢٠٠٧ و٢٠٠٨، والاختبار الرسمي في اثني عشر عامًا، وبرزت وظيفته في تقييم العملية التعليمية للغة العربية على نطاق الدولة.

يضم الاختبار ١٥٥ سؤالًا تتوزع على سبعة عناصر، ليُغطي بذلك مجموعة واسعة من الموضوعات، وتتصف الأسئلة الواردة فيه بالنموذجية، وتشمل القواعد الصرفية والنحوية والإملائية الأساسية، وكيفية بناء الجملة والاستخدام الملائم للألفاظ، وهذا يعني أن الاختبار يقيّم الكفاءة اللغوية عند المتحَنين، وقدرتهم على التطبيق اللغوي المتكامل، الأمر الذي يكسبه ميزات العلمية والموضوعية والقياسية، ويجعله أكثر الاختبارات شمولية وفاعلية ومرجعية في إظهار مستوى الطلبة الحقيقي، ونوعية التدريس. والمعروف أن الهدف النهائي من الاختبار هو خدمة العملية التعليمية والارتقاء بجودتها على نحو مطّرد. وبعد دخول القرن الجديد شهد تخصص اللغة العربية في الجامعات الصينية توسعًا سريعًا، ومع ذلك لم يشهد مستوى التعليم ما يماثله من التطور، بل على العكس من ذلك شَهِدَ تراجعًا ملحوظًا في المهارات اللغوية الأساسية عند الطلاب، وقدرتهم على التطبيق اللغوي. وقد وضع الاختبار معيارًا وطنيًّا موحدا لقياس نوعية التدريس

في المرحلة الأساسية، وذلك لأجل حث تخصص اللغة العربية على الاقتراب من هذا المعيار أو تلبيته أو حتى تجاوزه. وفي الوقت نفسه، يهدف الاختبار إلى تحفيز تخصص اللغة العربية في مختلف الجامعات لقياس مستوى التدريس الخاص به، وتحديد نقاط القوة والضعف عنده، حتى ترتقي نوعية العملية التعليمية على نحو متواصل. واستنادًا إلى ذلك تجمع الأوساط التعليمية الصينية للغة العربية على أن الاختبار العام أدى وما يزال يؤدي دورًا نشطًا في تثبيت المعايير التعليمية، وضبط المناهج الدراسية، وتأليف الكتب المنهجية، وتعزيز إصلاح العملية التعليمية، وتحفيز المدرسين على مواصلة التعلم والرقي بالذات، والعمل على تغييرها نحو الأفضل من حيث المستوى اللغوي والأكاديمي والفن التعليمي.

وفي الوقت نفسه يعد الاختبار العام بمثابة «مشير الريح»، يرسم للطلبة هدفًا دراسيًّا محددًا. فقد نص «المنهج» على أن هدف التعليم في المرحلة الأساسية لتخصص اللغة العربية في الجامعات الصينية هو «جعل الطلاب يحيطون بأنواع المعارف الأساسية في اللغة العربية؛ وجعلهم يستوعبون مهارات التطبيق اللغوي من خلال التدريب الواعي والشامل؛ وجعلهم يلمّون إلمامًا أوليًّا بالبلدان العربية ثقافتها وتاريخها وأوضاعها؛ وجعلهم يتقنون فن الحوار والتواصل باللغة العربية على نحو أولي، وجعلهم يجيدون طرق التعلم السليمة، وتعويدهم على العادات الدراسية الجيدة، لإرساء أساس لغوي متين لدراستهم في المرحلة الدراسية المتقدمة.» (المنهج، ٢٠٠٠) وعلى ضوء ذلك تمت صياغة منهاج الاختبار الذي يعد تجسيدًا حيًّا للتوجيهات الموضحة في «المنهج» بشأن متطلبات التدريس وأهداف التعلم. ويعد الاختبار العام صورة مبيّنة لأهداف التعليم والتعلم، لذا يستحسن للطلبة عَدّ الاختبار كأحد أهدافهم الدراسية، والعمل على استيعاب محتواه والحرص على تحويله إلى قدرة لغوية حقيقية، والسعي جاهدين لاستيفاء متطلباته بشكل كامل في الدراسة اليومية، فيرتفع المستوى اللغوي شفويًّا وكتابيًّا، وتتحقق التنمية الشاملة لهم.

هذه الطبعة إهداء من المجمع
ولا يُسمح بنشرها ورقياً أو تداولها تجارياً

وإلى جانب ذلك يُعَد الاختبار مرجعًا هامًا للوزارات والشركات وأصحاب العمل في توظيف الخريجين، واختيار الأكفاء منهم. والاختبار بوصفه اختبارًا وطنيًّا وحيدًا يشارك فيه كل المتخصصين الجامعيين في اللغة العربية، ترتفع مرجعيته وتأثيراته الاجتماعية مع مرور السنوات. وتضم شهادات الاختبار التي تُمنَح من قبل مجلس التوجيه، ثلاث تقديرات، هي «امتياز» و«جيد جدا» و«مقبول»، تعكس بصورة حية ومباشرة اعتراف الدولة بمستوى الطلبة في اللغة العربية، فتعترف بقيمتها و«محتواها الذهبي» الأوساط الاجتماعية الصينية. وبعبارة أخرى أن أداء الطالب في الاختبار العام يؤثر تأثيرًا مباشرًا في الفرص المتاحة له في التوظيف وإكمال الدراسات العليا، إذ إن كل الوزارات والأجهزة والوكالات والشركات الحكومية تطلب من الخريجين «شهادات اختبار المستوى الرابع» حتى يتم تعيينهم في وظائف حكومية رسمية.

5 - المأمول من اختبار المستوى الرابع لتخصص اللغة العربية في الجامعات الصينية

لقد أثبت التطبيق العملي والتجارب والعبر المستفادة من إجراء الاختبار العام في الأعوام الاثني عشر الماضية أن الاختبار يساعد الممتحَن على إيجاد الفجوة بين التقييم الذاتي والتقييم المعياري، وإيجاد أوجه القصور في عملية التعلم، فيضع لنفسه وجهة أو فكرة للهدف الذي سيسعى للوصول إليه في تعلم اللغة العربية في الفترات الزمنية القصيرة والمتوسطة والبعيدة المدى، ثم يقوم بتحديد الخطوات المتتابعة بأكبر قدر ممكن من التفاصيل والجلاء، حتى يمكن تحقيق الهدف على أرض الواقع، وبدون أن يستهلك منه سدى الجهد والوقت والطاقة. أما بالنسبة إلى المدرسين أو مسؤولي تخصص اللغة العربية فيمكن أن يُعَد الاختبار العام وسيلة هامة لتكوين تصور واضح وعميق عما قد استوعبه الطلبة، فيقدرون على تطبيق خططهم المستقبلية على ضوء ما يمتلكونه من معارف ومهارات نتيجة لتعلمهم؛

وكذلك يمكن عَدّه أداة ناجحة لجمع المعلومات حول مدى نجاح طرق التدريس المستعملة، ومدى نجاح المصادر التعليمية المساندة المستخدمة في مساعدة الطلبة على إتقان ما درسوه، ومدى نجاح طرق الإدارة الصفية المستخدمة في توفير بيئة تعلم جاذبة وداعمة لهم، وهكذا يقف المدرسون وقوفًا علميًّا على نقاط القوة والضعف عند الطلبة، وأوجه القصور في تنفيذ أهداف التعليم، ومن ثم يدركون ما هي العناصر التي تحتاج إلى تطويرها بغرض رفع مستوى الأداء التعليمي. أما بالنسبة إلى الجهات المسؤولة عن تنفيذ برنامج توظيف الطلاب الجامعيين، فتساعدها درجاتهم في الاختبار العام على التعرف على مستواهم اللغوي الأساسي، فتنجح في اختيار الأكفاء المتميزين والأجدر لشغل الوظيفة.

وفي عصر الشبكات الرقمية والبيانات الضخمة، شهدت أنشطة الإنسان وتفاعله مع العالم والبيئة المحيطة به متغيرات جبارة، كما شهدت خصائص التعلم، تلك العملية الحيوية الديناميكية عند الطلاب، تغيرات كبيرة أيضًا. وبسبب أن طلبة اليوم من الجيل الذي نشأ متزامنًا مع تطور تكنولوجيا الحاسوب والشابكة والهواتف النقالة الذكية وغيرها من المنتجات الرقمية، تتجلى ظروف نشأتهم بصورة واضحة في أنماط تفاعلهم مع البيئة المادية والاجتماعية والمعرفية. وتماشيًا مع ذلك تغيرت متطلبات المجتمع تجاه كفاءات الموارد البشرية المتخصصة في اللغات الأجنبية، فباتت الكفاءات التي تتقن اللغة الأجنبية فقط أمام تحديات العصر، أما الكفاءات التي تدمج بين إتقان اللغات الأجنبية وجوانب من العلوم الأخرى فأصبحت محط ترحيب. وفي ظل هذا الوضع الجديد، بدأ تخصص اللغة العربية في الصين يبذل جهدًا في تكوين أكفاء ذوي مستوى متين في اللغة العربية، ووعي قوي بمبادئ التواصل بين الثقافات، واطلاع واسع على آخر التطورات في القضية الشرق أوسطية والقضايا العالمية، ومستوى معين من المعارف المهنية، وقدرة تطبيقية قوية، ورصيد ثقافي جيد، ووعي إبداعي قيم، وقوة تنافسية شاملة. ومن أجل تحقيق هذا الهدف، تم وضع

«المعيار الوطني لتعليم اللغة العربية في الجامعات والمعاهد العالية الصينية» (يشار إليه فيما يلي باختصار «المعيار الوطني»)، الذي يسلط الضوء على توسيع النطاق والمجالات في التعليم الاختصاصي، وإبراز الخصائص الإنسانية والعلمية في عملية إعداد الكفاءات المتخصصة في اللغة العربية، وقدرتها على التأقلم مع المتغيرات الاجتماعية، بحيث ينص على أن تخصص اللغة العربية في الجامعات الصينية «يحمل على كاهله رسالة إعداد كفاءات ذات رؤية عالمية للخصائص الإنسانية، وتجيد اللغة العربية وآدابها، وتستوعب المعارف الثقافية العربية والإسلامية، وتتمتع بالقدرة على التطبيق اللغوي المتكامل، والقدرة على التواصل بين الثقافات، والتفكير الجدلي والابتكار، وإصدار الأحكام المعقولة والصائبة على الأحداث والأوضاع، والقدرة على الترجمة وإدارة الشؤون الخارجية وممارسة التدريس، والقيام بالبحوث الأكاديمية بشكل أولي.» (المعيار الوطني، ٢٠١٦) ويعد المعيار الوطني وثيقة منهجية تمت صياغتها تكيفًا مع خصائص التعليم والتعلم في عصر الإنترنت، وتماشيًا مع متطلبات الإصلاح التعليمي والخطة التعليمية التطويرية الطويلة الأمد على مستوى الصين، وأيضًا من أجل تعزيز إصلاح التعليم الجامعي، وتحسين الجودة في إعداد كفاءات متميزة قادرة على المنافسة العالمية، والمساهمة بشكل فعال في مختلف المجالات لتخصص اللغة العربية. والاختبار العام بوصفه وسيلة تقييم وطنية وحيدة وبمشاركة كل المتخصصين للغة العربية، لا بد أن يساير توجهات التنمية الاجتماعية والإصلاح التعليمي، ويطبق بشكل جدي روح «المعيار الوطني» من خلال مواصلة الإصلاح والابتكار، حتى ينجح في تلبية متطلبات التقييم في نوعية التعليم للكفاءات المتخصصة في اللغة العربية، وبالتالي الارتقاء بمستوى التعليم لتخصص اللغة العربية بالصين إلى مستوى أعلى.

الأول: أهداف الإصلاح للاختبار العام

١. مواصلة تعزيز دور التغذية الراجعة الإيجابية للاختبار العام

فقد أشار العالمان المعروفان في مجال الاختبار اللغوي ألدرسن جاي سي (Alderson J C) ووول دي (Wall D) إلى أن «الاختبار يمكن أن يؤثر في العملية التعليمية في خمسة عشر منحى، منها محتوى التعليم، وموقف المتعلم من التعلم، وطريقته ومستواه وعمقه في التعلم، ويختلف دور الاختبارات في التغذية الراجعة للعملية التعليمية باختلاف أهمية الاختبارات ذاتها.» (Alderson J C & Wall D. 1993) والاختبار العام بسبب أهميته الكبرى التي ذكرناها سابقًا يُعَد عاملًا مهمًّا ومؤثرًا في العملية التعليمية، ويبدو دوره في التغذية الراجعة للعملية التعليمية واضحًا جليًّا.

وذهب هيوز أي (Hughes A.) العالم الذائع الصيت في الاختبار اللغوي إلى أن «التغذية الراجعة للاختبار قد تكون إيجابية، فيؤدي الاختبار دورًا تحفيزيًّا إيجابيًّا للعملية التعليمية؛ وقد تكون سلبية، فيؤدي الاختبار دورًا سلبيًّا يحول دون تحقيق الهدف المصاغ للعملية التعليمية.» (Hughes A. 2003) والاختبار العام للغة العربية لا يخرج عن هذه النواميس، وهو بشقيه الإيجابي والسلبي، قد يؤثر تأثيرًا إيجابيًّا في تعليم اللغة العربية في الصين، وقد يؤثر تأثيرا سلبيًّا عليه. وتتجلى آثاره السلبية في جانبين: أولًا، بسبب صعوبة التشغيل، حيث لا يشمل الاختبار العام نوع أسئلة الاختبار الشفوي، رغم أن أهمية قدرة الطلبة على التعبير الشفوي أمر لا يقبل الجدال. وبسبب غياب هذا النوع من الاختبار قد يصرف المدرسون والمتعلمون جهودهم وطاقاتهم للعناصر الواردة في الاختبار، مما يقلل من اهتمامهم وعنايتهم بتطوير قدرتهم على التعبير الشفوي، علمًا منا أن استيعابها يتطلب منهم التدريب الشاق والمتواصل حتى يمتلكوا القدرة على نقل الفكرة ارتجالا، وبصورة واضحة ومتسلسلة. ثانيًا، الانخراط المفرط

هذه الطبعة إهـداء من المجمع
ولا يُسمح بنشرها ورقياً أو تداولها تجارياً

في بحر الأسئلة لأجل تسجيل معدل نجاح أعلى في الاختبار. فقد تصرف بعض تخصصات اللغة العربية الجهود والطاقات في تحليل أنماط ومحتويات الاختبار أكثر من اللازم، أو حتى تجعل التعليم يدور أساسًا حول الاختبار، فهذا خطأ فادح، حيث إنه يجعل المتعلمين يتجاهلون تنمية المهارات والقدرة على التطبيق اللغوي، ويغضون نظرهم عن التنمية الشاملة، فهذا يجانب الصواب في التعليم، وينحرف عن الهدف المرسوم لإقامة الاختبار العام. ومع ذلك لا ننفي أن المذاكرة والمراجعة بالقدر المناسب مفيدتان، إذ إن في الإعادة إفادة.

٢. ينبغي للاختبار العام تسليط الضوء على قياس القدرات السبع المحددة في المعيار الوطني

لقد مرت الاختبارات اللغوية بمرحلة ما قبل الاختبار العلمي التي اتصف فيها الاختبار بحصر أسئلته على ما تم تدريسه فقط؛ ثم بمرحلة البنيوية، أي مرحلة القياس النفسي التي اتصف فيها الاختبار بالمعيارية والموضوعية؛ مرورًا بمرحلة اللسانيات الاجتماعية التي سلط الاختبار فيها الضوء على قياس القدرة اللغوية الخالصة من خلال طريقة الاختبار الشامل؛ ومرحلة قياس الكفاءة التواصلية التي استهدف فيها الاختبار قياس المعرفة الضمنية للغة والقدرة على استخدامها بفعالية، وكان هذا النوع من الاختبار شائعًا في ثمانينيات القرن العشرين؛ أما بعد التسعينيات فقد أصبح القياس اللغوي المستند إلى المهام نقطة ساخنة جديدة في البحث والدراسة. (هان، ٢٠٠٣) ومن خلال استعراض تاريخ تطور الاختبارات اللغوية، يمكن ملاحظة أن الاختبار اللغوي يُعَدّ تجسيدًا حيًّا للنظريات اللغوية، بينما يكون الاختبار العام عندنا يمزج في وضع أسئلته بين المكوّنات المعقولة والعلمية في مختلف المراحل، منها مرحلة البنيوية، ومرحلة اللسانيات الاجتماعية، ومرحلة قياس الكفاءة التواصلية، ويتخذ بشكل أساسي مبدأ قياس قدرة الطلبة اللغوية كهدف رئيس له.

اختبارات اللغة
Journal of Language Testing

وينص المعيار الوطني على أن المتخصصين في اللغة العربية وآدابها يجب أن يسلحوا أنفسهم بسبعة أنواع من القدرات، وهي القدرة على التطبيق اللغوي المتكامل، والقدرة على التذوق الأدبي، والقدرة على إصدار الأحكام المعقولة والصائبة على الأحداث والأوضاع الاجتماعية، والقدرة على التواصل بين الثقافات، والقدرة على التفكير الجدلي والابتكار، والقدرة على التعلم الذاتي المستقلّ، والقدرة على التطبيق والممارسة. (المعيار الوطني، ٢٠٠٦) ومن الملاحظ أن المعيار الوطني يتبع المنهج العلمي في تحديد القدرات التي يجب على الطلبة المؤهلين استيعابها، ويؤدي دورًا توجيهيًّا وطليعيًّا في تعزيز العملية التعليمية وزيادة فاعلية الاختبار. ويلزم للاختبار العام العمل على تحسين فاعلية العناصر التي تم تحديدها للاختبار، والسعي لتطوير أنواع جديدة من الأسئلة والعناصر لتعزيز علميته وشموليته، حتى يسجل نجاحًا أكبر في توجيه تخصص اللغة العربية في مختلف الجامعات لتنمية القدرات السبع عند الطلبة، وبالتالي يساهم بنصيبه في رفع جودة التعليم الجامعي، وتمكين الطلبة من التنمية الشاملة والمتوازنة والمتكاملة لقدراتهم. فقد أثبت الاختبار العام من خلال اثني عشر عامًا من التجربة أن معظم الأسئلة الواردة فيه نافعة، ولها فاعلية كبيرة في قياس القدرة اللغوية عند المتحنين، وبذلك حقق الاختبار الهدف المتوقع منه. ولكن مع تزايد ما توليه الاختبارات اللغوية من الاهتمام بقياس القدرة الشاملة عند المتحنين، صارت التصاميم لمهام الاختبار أكثر شمولية، بعد أن كانت تستهدف قياس قدرة منفصلة محددة عندهم فقط. وتسايرًا مع تطور المفهوم في علم الاختبارات اللغوية، يلزم للاختبار العام مواكبة ركب التطور، والعمل على إدخال إصلاحات شاملة على نفسه.

الثاني: سبل رفع كفاءة الاختبار العام

بتأثير النظريات المعاصرة بشأن التربية والتعليم أخذت الاختبارات اللغوية منحى مغايرًا لما كانت عليه، وصار تخصص اللغة العربية يواكب التطور الحضري والتقدم الأكاديمي حتى يحقق نتائج تعليمية أعظم، فأصبح الاختبار

هذه الطبعة إهداء من المجمع
ولا يُسمح بنشرها ورقياً أو تداولها تجارياً

يُعنى بقياس كل العملية التعليمية من أجل الحكم على مستوى تحصيل الطلبة واستيعابهم وفهمهم لما يجب عليهم استيعابه، كما يُعنى بالكشف عن مدى فاعلية التدريس والمنهج التعليمي والكتب المدرسية.

١. زيادة أسئلة الاختبار الشفوي

ومن أجل أن ينجح الاختبار العام في قياس مدى قدرة الطلبة على التطبيق اللغوي المتكامل، وقدرتهم على التواصل عبر الثقافات، نقترح زيادة أسئلة الاختبار الشفوي. وعند التصميم والصياغة لمثل هذا النوع من الأسئلة، يجب مراعاة التفاعلية لصيغ الأسئلة، ومراعاة كيفية ارتباط محتوى الأسئلة بالحياة الواقعية والمشاهد الحقيقية في التواصل بين الثقافات، حتى لا يشعر الطلبة بعدم جدوى ما تعلموه، أو يشعروا بأن ما درسوه يتصف بالتجريد، أو الجفاف، أو البعد عن المواقف الحياتية الحقيقية. كما يمكن دمج طريقة قياس قدرة الطلبة على التعبير الشفوي مع طريقة قياس قدرتهم على القراءة الجهرية وفهم المسموع، علمًا منا أن «القراءة الجهرية تعتمد على ثلاثة عناصر أساسية هي: رؤية الرمز المكتوب بالعين، ونشاط الذهن في إدراك معنى الرمز، والتلفظ بصوت مسموع يعبر عن دلالة الرمز أو الحرف.» (الكحلوت، ٢٠٢١) ومن خلال القراءة الجهرية، يمكن قياس مدى إجادة الطلبة للنطق وتمثيل المعاني والطلاقة في التعبير وصحة الفهم للمقروء؛ ومن خلال إجابة الطلبة عن الأسئلة المطروحة باللغة العربية يمكن قياس ردود أفعالهم واطلاعهم ومعارفهم، وقدرتهم على فهم المسموع وتنظيم المعلومات وترتيب الأفكار والاسترسال بالحديث.

٢. إصلاح صيغ السؤال المتعلق بالتعبير الكتابي حتى تكون أكثر تنوعًا.

ومن المتفق عليه أن التعبير الكتابي له أهمية خاصة في التعليم اللغوي في تجسيد مكتسبات الطلبة من خلال وصفهم للمشاهد والأحداث والشخصيات، وتعبيرهم عن الأفكار والمشاعر والأحاسيس والآراء، وله دور حيوي في توسيع خيالهم وتنمية إبداعهم. ولذلك خصص الاختبار العام خمس عشرة درجة ليمتحن قدرة الطلبة

على التعبير الكتابي. رغم ذلك نجد أن طريقة صياغة السؤال المتعلق بالتعبير الكتابي جامدة، وهو اختيار موضوع من الموضوعين لكتابة إنشاء، وكتابة مقال تطبيقي حسب المطلوب. وإلى جانب ذلك يمكن دمج التعبير الكتابي مع فهم المقروء، أو إكمال مقالة لها بداية محددة، أو وصف صور ورسومات، لزيادة مرونة الأسئلة وتحفيز الطلبة على إطلاق العنان لمخيلاتهم وعرض إبداعاتهم.

٣. زيادة نسبة فهم المقروء في الاختبار

يعد فهم المقروء من أساسيات المهارات اللغوية، وله علاقة مباشرة بنجاح الطالب في حياته العملية والأكاديمية في المستقبل، نظرًا لأن الفهم القرائي هو نتيجة للعديد من النشاطات العقلية التي تتم تلقائيًّا ودون وعي منا، منها فهم المعنى الذي تنقله المقالة من خلال التحليل والاستقراء، والتنبؤ بمجريات الأمور والأحداث على أساس المنطق، وتفعيل الخبرات الحياتية أو المعرفة القرائية السابقة، وتحديد غرض الكاتب من كتابة المقالة، وذلك إلى جانب أن يستوعب الطالب كمية مطلوبة من عدد الكلمات، وقواعد اللغة العربية في النحو والصرف والبلاغة، والقدرة على تحديد معاني الألفاظ والجمل على ضوء أدلة السياق. وعليه فإن عنصر فهم المقروء في الاختبار العام هو نوع من الأسئلة ذات الشمولية في قياس قدرات الطالب على الفهم والتقدير والاستنتاج والتلخيص والتذوق. لذا نقترح زيادة نسبة الأسئلة الخاصة بفهم المقروء استهدافًا لتوجيه المدرس والطالب إلى إيلاء اهتمام أكبر للقراءة داخل النشاطات التعليمية الصفية أو خارجها، والعمل على تعزيز استيعاب الطالب لما درسه في الفصول الدراسية، وتوسيع المفردات، وتنمية الذوق اللغوي، ورفع القدرة على التطبيق اللغوي من خلال القراءة والمطالعة. وفي تصميم الأسئلة يمكن تسليط الضوء على فهم واستيعاب الطالب للفكرة الرئيسة للمقالة، وقدرته على التفكير الجدلي، مع إيلاء عناية لقدرته على التعبير عن رأيه الخاص في ما قرأه.

هذه الطبعة إهداء من المجمع
ولا يُسمح بنشرها ورقياً أو تداولها تجارياً

٤. زيادة عدد الأسئلة المتعلقة بالأوضاع في البلدان العربية والثقافة العربية الإسلامية

إن اللغة العربية وعاء للثقافة العربية الإسلامية، وهي الأداة المثلى للتعرف على تاريخ البلدان العربية والأوضاع الراهنة فيها اقتصاديًا وسياسيًا واجتماعيًا. ومن غير المعقول أن نركز جهودنا على تدريس ودراسة اللغة العربية بإغفال المكوّن الثقافي لها. لذا يلزم للاختبار العام زيادة عدد الأسئلة المتعلقة بأوضاع البلدان العربية والثقافة العربية الإسلامية، حتى ينجح في توجيه الطلبة لصرف مزيد من العناية بالصفات الإنسانية والثقافية للغة العربية، وبالتالي يتعمق فهمهم للخواص الجوهرية للغة، فيتقنون الاستخدام اللائق والملائم لها، ويرتقي المستوى اللغوي والفكري والثقافي عندهم.

٥. تطوير اختبار المستوى الثامن لإقامة منظومة اختبارية متكاملة

قد مر اثنا عشر عامًا على إقامة اختبار المستوى الرابع، وتراكمت عند فريق الخبراء خبرات لازمة لتطوير اختبار المستوى الثامن. ولأجل إنجاز الهدف المرسوم لا بد للخبراء من التعريف الأكثر علمية بالقدرات المطلوب اختبارها في المستويين الرابع والثامن كل على حدة، حتى يمكن للاختبار أن يؤدي دوره في ضبط العملية التعليمية في تخصص اللغة العربية في الصين. وعلى سبيل المثال، يمكن لاختبار المستوى الثامن أن يقوم بتصميم أكبر عدد من الأسئلة المقالية وصياغتها؛ ليقيس بفاعلية أكبر القدرات السبع المدرجة في المعيار الوطني ويقيّمها. ومن المعلوم أن ظاهرة التراخي في دراسة اللغة العربية منتشرة بين الطلبة بعد تقدمهم إلى اختبار المستوى الرابع، والأسوأ من ذلك أن المقررات التخصصية تقل بنسبة كبيرة بعد انتقالهم إلى السنة الدراسية الجامعية الرابعة، ناهيك عن ما يشغلهم عن الانهماك في دراسة اللغة العربية من التطبيق العملي والبحث عن فرص التوظيف. ومن هذا المنطلق نجد أن إنشاء اختبار المستوى الثامن سيفرض ضغوطًا على الطلبة، وفي الوقت نفسه يشكل قوة تحفيزية عندهم.

اختبارات اللغة
Journal of Language Testing

الخاتمة

يعد اختبار المستوى الرابع أكثر أساليب التقييم فاعلية لقياس الكفايات اللغوية الأساسية عند الطلبة بعد إكمالهم المرحلة الدراسية الجامعية الأساسية في تخصص اللغة العربية بالصين، ولعب وما يزال يلعب دورًا إيجابيًّا في دفع الجامعات لإيلاء اهتمام لرفع المهارات اللغوية الأساسية عند الطلبة، وضبط التخطيط العلمي للمقررات، فلا يحل محله أي اختبار آخر. وبما أن الغرض الأساسي من إقامة الاختبار العام هو قياس مدى استيعاب الطلبة للمعارف والمهارات اللغوية، ومدى اتجاههم الإيجابي نحو تعلم اللغة العربية، ثم تقديم تغذية راجعة واضحة، يمكن لتخصص اللغة العربية الاستفادة منها عقب تحليل أداء الطلبة في كل عنصر من عناصر الاختبار، لا بد للجنة الاختبار من اتباع العلمية والمصداقية والموثوقية في تصميم الأسئلة وصياغتها، ومن إدخال إصلاحات على الاختبار؛ ليتفق مع ما تم النص عليه من المتطلبات المعرفية والمهارية والكفاءات الشاملة في المعيار الوطني، فيسعى تخصص اللغة العربية جاهدًا لتطوير طرق التدريس، وإثارة الدافعية الذاتية عند الطلبة، ليشاركوا المدرسين في إنجاز الهدف التعليمي، ورفع المستوى اللغوي والأكاديمي، وبالتالي تحقيق النمو الشامل.

المراجع:

- شيو، دونغ هوي، (٢٠٠٤). "تأملات في استكمال منظومة اختبارات اللغة الروسية في الجامعات والمعاهد العالية بالصين". مجلة "تعليم اللغة الروسية في الصين" (٣).
- الكحلوت، عماد حنون، والكحلوت ريم شحدة، (٢٠٢١). مدى تحقيق مشروع تحدي القراءة لأهدافه لدى طلبة وكالة الغوث الدولية في فلسطين "دراسة تقيمية". مجلة الدراسات والبحوث التربوية(١).
- ليو رون تشينغ، هان باو تشنغ، (٢٠٠١).الاختبارات اللغوية وطرقها. بكين: دار تعليم اللغات الأجنبية والبحث فيها.
- لجنة إعداد المعيار الوطني لتعليم اللغة العربية في الصين، (٢٠١٦). المعيار الوطني لتعليم اللغة العربية في الجامعات والمعاهد العالية الصينية. لم يتم إصداره.
- لجنة إعداد منهج تعليم اللغة العربية في الصين، (٢٠٠٠). منهج تعليم اللغة العربية في الجامعات الصينية. بكين: دار نشر جامعة بكين.
- هان، باو تشنغ، (٢٠٠٣). آخر التطورات في الاختبار اللغوي: الاختبار اللغوي المستند إلى المهام. مجلة "تعليم اللغات الأجنبية والبحث فيها" (٥).
- وو، تسون مين، (٢٠٠٤). اختبارات اللغة الإنجليزية -نظريا وتطبيقيا. بكين: دار تعليم اللغات الأجنبية والبحث فيها.
- Alderson J C & Wall D. (1993). Does washback exist?Applied Linguistics.
- Hughes A. (2003). Testing for Language Teachers. Cambridge: Cambridge University Press.

اختبارات اللغة
Journal of Language Testing

هذه الطبعة إهـداء من المجمـع
ولا يُسمح بنشرها ورقياً أو تداولها تجارياً

دراسةٌ مقارنة بين توجيهاتِ المجلس الأمريكي لتعليم اللغات الأجنبية والإطارِ الأوربيّ المرجعيّ المشترك للُّغات

د.إسلام يسري الحدقي[1]

مستخلص

ظهرت أُطر الكفاءة اللُّغوية في اللُّغات الأجنبية في الولايات المتحدة الأمريكية في منتصف القرن العشرين بهدف الاعتماد على تقييم موضوعيّ للمستويات اللُّغوية لدى متعلمي اللُّغات الأجنبية من المعنيين بها في الإدارة الأمريكية، ومنذ ذلك الوقت بدأت الأُطر تنشأ وتتطور فظهرت أُطر حسب جمهور اللُّغات وحاجاتهم؛ منها ما يُوَجّه لتقييم الموظفين الحكوميين، ومنها ما يقيّم العاملين في المجال العسكريّ، ومنها ما يناسب المجال الأكاديميّ، والبعض ليلائم المتعلمين الصغار، وآخر يُقيّم المهاجرين، وغيره يقيس أداء متعلمي اللُّغة لأغراض خاصّة... إلخ. ومع انتشار برامج إعداد معلمي العربية للناطقين بغيرها حول العالم يُطْرَحُ سؤال باستمرار: أي الأُطر أفضل؟ بل إنَّ البعض يميل لإطارٍ على حساب آخرَ دون براهين علمية واضحة. ومما لاحظه الباحث في الفعاليات العلمية المختلفة أنَّ هناك الكثيرَ من المعلومات الغائبة عن الأُطر لدى الكثيرين، وهو ما يجعل الصورة الكُلّية عن هذه الأُطر غير دقيقة، وبالتبعية لا يوجد منطق للتحزُّب لإطار بعينه. كما أنَّ هناك جوانب مختلفة للأُطر لكي نحكم على مناسبتها لسياق مُعيّن

(١) أستاذ مساعد ومنسق الاختبارات العربية بجامعة السلطان محمد الفاتح الوقفية

يتجاوز حدود عدد المستويات وواصفاتها؛ لذا حاولت هذه الدراسة الإجابة عن سؤال: ما مواضع الاتفاق والاختلاف بين توجيهات المجلس الأمريكي لتعليم اللُّغات الأجنبية والإطار الأوربي المرجعي المشترك للغات؟

توصلت الدراسة إلى نتائج من أهمها: إن الإطارين اتفقا في مساحات أبرزها الاعتماد على أُطر سابقة لهما، وخضوعهما للتجريب قبل الإصدار النهائي، ونشرهما إصدارين، واتفاقهما في استعمالهما في التقييم وتقديم الدعم للمستخدمين، واختلفا في سَعة استخدامات الإطار الأوربيّ عن توجيهات المجلس الأمريكيّ، وتميزت الأخيرة بصغر حجمها مقارنةً بمقاييس الإطار الأوربيّ، ووجود إضافات خاصّة باللُّغة العربية داعمة لواصفاتها، بالإضافة إلى توفُّر اختبارات كفاءة معتمدة وَفق توجيهاتها، كما جاءت أقسام الإطار الأوربي وواصفاته أكثر اتساعًا من توجيهات أكتفل فضلًا عن حداثة أقسامها مثل تناول التفاعُل عبر الإنترنت. ويُتوقَّع أن يُستفاد بهذه المقارنة في مساعدة المعنِيين بتعلُّم اللُّغات على اختيار الإطار المناسب لطبيعة جمهورهم وبحسب حاجاتهم اللُّغوية ووفقًا للبيئة التي يعملون فيها، بالإضافة إلى مساعدة المعنيين بفكرة الإطار المرجعي العربيّ على التخطيط لهذا المشروع واتخاذ القرارات في ضوء تجارب السابقين.

الكلمات المفتاحية:
الكفاءة اللُّغوية، الأُطر المرجعية، توجيهات أكتفل، الإطار الأوربي.

هذه الطبعة إهـداء من المجمـع
ولا يُسمح بنشرها ورقياً أو تداولها تجارياً

Abstract:

The frameworks of foreign language proficiency emerged in the United States in the mid-twentieth century with the aim of relying on an objective evaluation of language levels among foreign language learners who are relevant to the American administration. Since then, frameworks have begun to emerge and develop according to audiences and their needs. Some are directed towards evaluating government employees, some evaluate military personnel, some are suitable for academic fields, some for young learners, and others evaluate the languages of immigrants and for specific purposes, etc. With the spread of Arabic teacher-training programs for non-native speakers around the world, a question is constantly raised: Which framework is better? Some even tend to favor one framework over another without clear scientific evidence. In various scientific events, the researcher noticed that there is a lot of missing information about frameworks among many people, which makes the overall picture of these frameworks inaccurate. Therefore, there is no logic to favor a particular framework. There are also different aspects of frameworks that need to be considered when judging their suitability for a specific context beyond the number of levels and descriptors. Therefore, this study attempted to answer the question: What are the areas of similarities and differences between the American Council on Teaching Foreign Languages guidelines (ACTFL) and the Common European Framework of Reference for Languages (CEFR)?

The study yielded several important findings, including agreement between the two frameworks on key areas, such as reliance on previous frameworks, testing prior to final release, two versions of publication until now, and agreement on their use in assessment and support for users. However, they differed in terms of the European framework's broader range of uses compared with the American Council's guidelines. The latter was characterized by its smaller size compared to CEFR and its inclusion of Arabic language-specific additions to its descriptors. In addition, it provides certified proficiency tests according to the guidelines. The sections and descriptors of CEFR were broader than those of ACTFL, with newer sections,

such as online interactions, being included. This comparison is expected to assist language teachers in selecting an appropriate framework for their audiences and linguistic needs based on their working environment. It will also aid those interested in developing an Arabic reference framework for planning this project and making decisions based on previous experiences.

Keywords:

Language proficiency, reference frameworks, ACTFL guidelines, CEFR

هذه الطبعة إهـداء من المجمـع
ولا يُسمح بنشرها ورقياً أو تداولها تجارياً

المقدمة

نشـأت الأُطـر المرجعيـة في بدايـة الخمسـينيات مـن القـرن العشـرين ضمـن مبـادرة لمعهـد الخدمـات الخارجيـة الأمريكيـة (FSI) التابـع للحكومـة الأمريكيـة؛ وهـو مؤسّسـة مَعنيّـة بتدريـب الدبلوماسـيين الأمريكيـين وموظفـي الخدمـات المتعلقـة بالأجانـب علـى اللُّغـات الأجنبيـة، وكان الهـدف الوقـوف علـى مسـتويات الكفـاءة اللُّغويـة في التحـدث لـدى هـؤلاء الموظفـين.

عقـب هـذه المبـادرة مبـادرات أُخـرى عملـت علـى تطويـر إطـار FSI نفسـه، ثـم بُنيـت أُطـر أخـرى تسـتهدف جمهـورًا جديـدًا، فظهـر إطـار حلـف شـمال الأطلسـي المعـروف باسـم STANG 6001 للعامليـن في المجـال العسـكري الـدولي، والإطـار الأُسـترالي المعـروف باسـم التقييـم الـدولي للكفـاءة اللُّغويـة في اللغـة الثانيـة ISLPR مسـتهدفًا المهاجريـن والطـلاب، وتوجيهـات المجلـس الأمريكـي لتعليـم اللُّغـات الأجنبيـة ACTFL الـتي وُجّهـت للمجالـين الأكاديـمي والمِهـني، والمعايـير اللُّغويـة الكنديـة CLB الموجَّـه للمهاجريـن أيضـا، وتلاهـم الإطـار الأوربي المرجعـي المشـترك للغـات CEFR

منهج الدراسة

اتبـع البحـث المنهـج الوصفـي المقـارن الـذي ظهـر في وصـف خصائـص الأُطـر عـبر المـرور بخطـوات أربعـة: تحديـد موضـوع المقارنـة، تحديـد متغـيرات المقارنـة، تفسـير بيانـات موضـوع المقارنـة، الوصـول إلى نتـائج المقارنـة.

مشكلة الدراسة

وجـود انطباعيـة في تقييـم أُطـر الكفـاءة اللُّغويـة؛ ممـا اسـتلزم الوقـوف علـى مواطـن الاتفـاق والاختـلاف بـين تلـك الأُطـر واسـتعمالاتها ومـدى موثوقيتهـا ومـا توفـره مـن مصـادر داعمـة.

سؤال الدراسة (موضوع المقارنة):

ما مواضع الاتفاق والاختلاف بين توجيهات المجلس الأمريكي لتعليم اللُّغات الأجنبية والإطار الأوربي المرجعي المُشتَرَك للُّغات؟

أهداف الدراسة

الوقوف على المُشترك والمختلف بين الإطارين موضع الدراسة ليساعد المعنيين بعملية تعليم اللُّغات على اختيار المناسب لظروف بيئتهم التعليمية وأهداف جمهورهم من عملية التعلُّم.

أهمية الدراسة

يُعَدُّ هذا البحث من البحوث العربية القليلة -في حدود علم الباحث- التي قارنت بين هذين الإطارين من منظور البحث الحالي والذي يهتم بالجوانب المختلفة للإطار وليس مجرد مستويات الأداء اللُّغوي، مع العلم أن هناك بحوثًا أُخرى غير عربية قارنت بين واصفات الأداء في توجيهات أكتفل والإطار الأوربي المرجعي من ناحية، والإطار الأوربي المرجعي والمعايير اللُّغوية الكندية من ناحية أخرى وعادلت بين مستوياتها.

كما يُتوقّع أن تفيد هذه المقارنة في تطوير الأفكار المتعلقة ببناء إطار مرجعيّ عربيّ موحَّد للُّغة العربية بوصفها لغة أجنبية/ثانية.

حدود الدراسة (متغيرات المقارنة):

من حيث الأُطر: توجيهات المجلس الأمريكي لتعليم اللُّغات الأجنبية، والإطار الأوربي المرجعي المشترك للغات.

من حيث عناصر المقارنة:

١. النشأة والجمهور.

٢. الأهداف.

٣. الأُسس النظرية.

٤. الإصدارات والتطوير.

٥. التقنين.

٦. سلم المستويات وأقسامه.

٧. الواصفات.

٨. الدعم الفني /الخدمات.

٩. الأثر.

١٠. النقد.

وهذه المقارنة لا تشمل مقارنة مستويات الأداء في كلِّ إطار.

١ - مفهوم الأطر المرجعية ووظائفها وخصائصها

١-١ مفهوم الإطار

يلاحظ الباحث في هذه المضمار أنَّ كلَّ إطار من أُطر الكفاءة اللُّغوية جاء باسم مختلف إلا أنَّ المفهوم كان واحدًا، فنجد (توجيهات أكتفل) و (الإطار الأوربي) و(المعايير الكندية) و(مقياس المائدة المستديرة بين الوكالات) و(التقييم الدولي للكفاءة)، ويذكر North (١٩٩٥) أنَّ من هذه الأسماء: مستويات الكفاءة، مقاييس الكفاءة، ودرجات الكفاءة، إرشادات، معايير، مستويات، مقاييس، مراحل،

موازين، درجات. كما أن تعريفاتها في مجال تعليم اللُّغات الأجنبية يختلف بحسب منظور الكاتب. (North, 2000).

إلا أنَّ الأشيع الآن في حقل البحث العلمي العربي مصطلح (الأُطر العالمية للكفاءة اللُّغوية).

وأُطر الكفاءة ببساطة عبارة عن سُلَّم مستويات لُغوية يقابل كلَّ مستوى منها وصف لما يستطيع أن يفعله الإنسان في كلِّ مستوى من هذه المستويات باللُّغة. أو كما تصورها Harsch & Malone (2020) أنها ذاتُ بُعدين: رأسي وأفقي، ويمكن أن نَعُدَّ الرأسيَّ للمستويات والأفقيَّ لما يستطيع الشخص فعله باللُّغة.

هذه المستويات قد تحمِل أسماء مثل (مبتدئ، متوسط، متقدم،..إلخ) كما هو الحال في توجيهات أكتفل، أو تحمِلُ حروفا مثل (A,B and C) مثل المُتبع في الإطار الأوربي المرجعيّ المشترك للُّغات، أو تحمل أرقامًا (معيار١، معيار٢، معيار٣، ..إلخ) مثل المعايير اللُّغوية الكندية، وهذا يُعَدُّ البُعد الأوَّل، أما البُعد الثاني ففيه «واصفات» تصف ما الذي يمكن أن يقوم به الشخص باللُّغة الثانية أو الأجنبية في كلّ مستوى من تلك المستويات.

١-٢ وظائف الأُطر المرجعية:

الأصل في نشأة الأُطر المرجعية للكفاءة اللُّغوية كان التقييم اللغوي؛ حيث ظهرت تلبيةً للحاجة إلى معيارية التقييم؛ فقد كان التقييم يقوم -قديما- على الانطباعية باستخدام عبارات مطاطة؛ مثل: «طلق في اللغة الفرنسية» و «ممتاز في الألمانية». (Herzog, Martha, n.d)، إلا أنها صارت فيما بعد مُوَجِّهَات للمناهج والنظم التعليمية بشكل عام، ويمكن تحديد بعض هذه الوظائف كما يلي:

- الاستفادة منها في بناء البرامج والمقرَّرات كما هو الحال في الإطار الأوربي المرجعي (تعليم-تعلُّم-تقييم).

- تُوفِّر «قواعد» يمكن الاحتكام إليها في عملية محاسبة المسئولين عن تعليم اللُّغات في المؤسسات المَعنيَّة، وهذا يمثّلُ نوعًا من الشَّفافية المطلوبة أمام أصحاب المصلحة كأولياء الأمور مثلا.
- تساعد على وجود لغة مشتَرَكة بين المعنيين بتعليم اللُّغات على جميع المستويات المحليّة والإقليميّة والدولية، فعلى سبيل المثال نقول: إنّ هذه الوظيفة تتطلّب شخصًا في مستوى كفاءة B2 مثلا؛ فيعرف المتخصّص ما الذي يستطيع أن يفعله هذا الشخص باللُّغة بالضبط، كما يعرِفُ الشخصُ المتقدِّم لوظيفةٍ ما القدراتِ اللُّغويةَ المطلوبةَ لها فور قراءته أنَّ المستوى المطلوب هو B2.
- تُسهِّل الأُطر المرجعية معادلة شهادات الطلاب اللُّغوية، بما يُسهّل من انتقالهم من دولة لأخرى أو من نظامٍ تعليميّ لآخر. انظر (North, 2000؛Piccardo, 2020).

١-٣ خصائص الأُطر المرجعية:

وحتى تؤدي هذه الأُطر وظائفها المهمّة السابقة فلابد أن تتصف بخصائص من أهمها: الشمول، والوضوح، والترابط، والمرونة، والقابلية للقياس. (محمود، ٢٠٢٠)، وهو ما يتكامل مع (الحدقي، ٢٠٢١) الذي يرى أنها يجب أن تتصف بالتفصيل، والتدرُّج، والدِّقة، وسهولة الاستعمال.

وتعمَد المؤسّسات المطوِّرة لهذه الأُطر إلى تحديثها بين الحين والآخر لكي تزيدَ من قوة خصائصِها حتى تؤدي وظيفتها بشكل فعَّال.

٢ - المقارنة بين توجيهات المجلس الأمريكي والإطار الأوربي المرجعي

وفقًا لـ Tschirner (2012) لم تُجرَ دراسات مقارنة رسمية بين توجيهات أكتفل والإطار الأوربي وإطار المائدة المستديرة إلى عام ٢٠١٠. وفي عام ٢٠١٠ عقدت فعاليات متعدِّدة بين المجلس الأمريكي والمعنيين بمؤسسات الاختبارات في أوربا لإجراء معادلة بين الأُطر سابقة الذكر، ونتيجة هذه اللقاءات صدرت معادلة بين مستويات أكتفل ومستويات الإطار الأوربي (Harsch & Malone, 2020)، وكان من ثمار هذه اللّقاءات تقرير «تنسيب تقديرات الإطار الأوربي إلى تقييمات أكتفل» (ACTFL, 2016). أو بمعنى أبسط معادلة مستويات الإطار الأوربي بمستويات أكتفل.

ركزت الأعمال السابقة على معادلة المستويات ببعضها، أو بمعنى آخر مقارنة واصفات الأداء بين مستويات الأُطر، أما هذا البحث فيركّز على مقارنة خصائص الأُطر نفسِها، وكي نقارن بين إطارين أو نختار إطارًا نعتمد عليه فلا يكفي أن نرى الواصفات والمستويات فحسب بل لابُدّ من دراسة أهداف الإطار والجمهور المستهدَف واستعمالاته والخدمات التي تقدّمها المؤسّسة المالكة للإطار ومدى مراعاة اللّغة العربية في إصدارته، وغيرها من العناصر التي سنراها في الفقرات التالية، فعلى سبيل المثال بعض الأُطر لا تنشر واصفاتها مجانًا على الشابكة بل تفرِضُ أن تشتريها مثل: التقييم الدولي للكفاءة اللُّغوية (الإطار الأسترالي) ممّا يشكِّل صعوبة في تداولها بين المعنيين، وبعض الأُطر لا تقدّم خدمة الاختبارات المعتمَدة مثل: الإطار الأوربي ومن ثمَّ فأيّ اختبار منسوب للإطار الأوربي فهو على مسئولية صاحبه وليس لزامًا أنه يعبّر بدقة عن مستويات الإطار، وبعض الأُطر لا توفّر مواد تدريبية مجانية لتأهيل المختبرين مثل: أكتفل، وعليه كي يحصل أحد على هذا التأهيل فعليه الالتحاق بدورات المجلس الأمريكي، والتي قد لا تُناسب تكلفتها مدرسي الدول النامية، وبناء على ما سبق فإنّ أمر اعتماد إطار كفاءة ما في مؤسسة تعليمية ليس يسيرًا.

١-٢ النشأة والجمهور:

جاء الإصدار الأوّل لتوجيهات المجلس الأمريكي لتعليم اللُّغات الأجنبية (١٩٨٦) نتيجة عدم مناسبة إطار المائدة المستديرة ILR للسياق الأكاديمي؛ حيث كان الطلاب يحتاجون أكثر من عام دون تجاوز المستوى الأول من مستوياته، وهو أمر له أثران: الأول: أنه غير قادر على تمييز مستويات الطلاب الذين درسوا لعام دراسي؛ حيث يصنفون كلّهم في المستوى الأول، والأثر الثاني: هو الأثر النفسي؛ حيثُ لا يشعر الطالب أنه تجاوز أيّ مستوى رغم طول فترة دراسته، ومن ثَمَّ كان من الضروري إعداد معيار جديد يناسب ظروف الدراسة الأكاديمية التي تقسم لفصول دراسية محدودة الزمن بدلًا من مقياس ILR الذي كان مُصمَّمًا لموظفي الحكومة الأمريكية. (الحدقي Liskin-Gasparro, 1984؛ Harsch & Malone, 2020؛ 2021). ونتيجة لذلك جاءت توجيهات أكتفل لتُستَخدم في النطاق الأكاديمي وأماكن العمل. (Actfl, 2012a)

أمّا عن الإطار الأوربي المرجعي فإن بيكاردو – واحدة من ثلاثة مؤلفين للمُجلد المصاحب– تصفه بأنه «عميق الجذور» حيث بُني على دراسات لمجلس أوربا بدأت في الستينات من القرن الماضي وقد تناولت وضع تصور لتحليل الحاجات وتحديد المستوى اللُغوي المطلوب للقيام بوظائف الحياة، وهو ما تَمَثّل في كتب مثل «مستوى العتبة Threshold» و«مستوى الاستقلال اللغوي Independence» و تجريب واصفات «يستطيع أن يفعل»(Piccardo, 2020). ، ويذكر المؤلف الرئيس للإطار الأوربي أنه في سبيل تطوير الإطار بَنى ما يمكن أن نسميه خزانة واصفات من ٣٠ إطارًا معروفًا في ذلك الوقت من بينها توجيهات أكتفل. (North, 2000)

طُوّر الإطار الأوربي المرجعي في الفترة بين (١٩٩٣–١٩٩٦) وقد نُشر الإصدار الأول للإطار المرجعي الأوربي عام ٢٠٠١ في اللغتين الإنجليزية والفرنسية. (North, 2007)، وأعقبهما ترجمات لعشرات اللُّغات منها العربية.

ومن الاسم الرسمي للإطار (تعلّم، تعليم، تقييم) نفهم أنَّ جمهوره الأساسي المعنيون بالتعليم وفي مقدمتهم الطلاب والمعلمون والمقيِّمون عبر المراحل التعليمية المختلفة، ونطاقه المؤسّسات التعليمية والجهات المعنية بالتعليم.

ونخلص إلى أنَّ كلا الإطارين اعتمد على أعمال السابقين، وجاء تلبيةً لحاجات المعنيين باللُّغات لنظام محدَّد ودقيق لتقدير الكفاءة اللُّغوية في اللغة الأجنبية أو الثانية، وكانت توجيهات أكتفل أقدم بحوالي ١٥عامًا من الإطار الأوربي المرجعي.

٢-٢ الأهداف.

التطبيق الأساسي لتوجيهات أكتفل هو تقييم القُدرة اللُّغوية الوظيفية؛ حيث تهدف إلى استخدامها للتقييم العام في الأوساط الأكاديمية وأماكن العمل. (أكتفل، ٢٠١٢)، وهنا يتضح أن تصميم التوجيهات الأساسي موجَّه لعملية الاختبارات، أمّا الإطار الأوربي المرجعي فإنه أرحب من ذلك؛ حيث صُمِّم ليكون أساسًا مشتركًا للمقرّرات اللُّغوية والمناهج والاختبارات والكتب الدراسية. (Council of Europe, 2001).

٢-٣ الأسس النظرية.

هناك ثلاثة ملامح أساسية لتوجيهات أكتفل، أنها تصرح بوضوح بعدم استنادها لنظرية بعينها أو طريقة تدريس أو منهاج تربوي، كما تقدم المهاراتِ اللغويةَ بشكلها المنفصل المشهور (استماع، تحدث، قراءة، كتابة)، والثالث أنها عَدَّت المعيار المثالي للأداء اللغوي هو «الناطق الأصلي جيد التعليم». (Actfl, 2012a). وهذا الأخير قد جرَّ عليها نقدًا كما سنذكر لاحقًا.

أما الإطار الأوربي فيظهر من مخططه للكفاية التواصلية اقترابه من نموذج القدرة اللُّغوية التواصلية (Bachman & Palmer (1996) ونموذج الكفاية التواصلية Celce-Murcia, Dörnyei, & Thurrell (1995). (North, 2020,10)

لم ينظر الإطار الأوربي للمهارات الأربعة بشكلها الموجود في الأُطر الأخرى كإطار المائدة المستديرة وتوجيهات أكتفل بالشكل المنفصل المشهور، ولكن نظر إليها على أنها أنشطة تواصلية تتضمن (الاستقبال، الإنتاج، التفاعل، الوساطة).

كما أن الإطار الأوربي اعتمد الناطق الأصلي معيارًا في إصداره الأول ٢٠٠١، وقد ظهر هذا في عدد من الواصفات مثل الواصفة التالية في مقياس الاستيعاب الشفهي/فهم المسموع «ليس لديه أي صعوبة من أي نوع من اللغة المنطوقة، سواء كانت حية أو مبثوثة، يتم تقديمها بسرعة مثل اللغة الأم." (مجلس أوروبا، ٢٠٢٠، ٢٩٣) إلا أنه عدَلَ عن ذلك في المجلد المصاحب؛ حيث رأى أن بعض الواصفات التي تشير إلى التوافق اللُّغوي أو عدمه مع المتحدثين الأصليين أثارت جدلًا منذ إطلاق الإصدار الأول للإطار عام ٢٠٠١. (السابق، ص ٣٨). لذلك تعامل مع متعلم اللغة بعَدِّه وسيطًا لغويًّا؛ أي يقوم بالمهام التواصلية دون مقارنته بالناطق الأصلي، وقد عُدِّلَتْ الواصفة في المثال السابق لتصبح «يستطيع أن يفهم بسهولة عمليًّا أي نوع من اللغة المنطوقة أو الإشارية، سواء كانت حية أو مبثوثة، يتم تقديمها بسرعة طبيعية».

(السابق، ص ٢٩٣؛ North, Piccardo, Goodier, Fasoglio, Margonis-Pasinetti, & Rüschoff, 2022)

٤-٢ الإصدارات والتطوير

إن الأُطر المرجعية انعكاس لأفكار مُطوريها عن تدرج عملية اكتساب اللُّغة وعن مكونات الكفاءة اللُّغوية نفسها، وتصدُق هذه الأفكار أو تكذب حسب نتائج تجريبها في ميدان الواقع، وكذلك حسب ما تظهره الأبحاث عن مكونات الكفاءة ومراحل اكتسابها. ويمكن تلخيص ذلك بالعبارة التي كُتِبَتْ في مقدمة توجيهات (Actfl, 1986) «إن عملية بناء توجيهات الكفاءة اللُّغوية عملية ديناميكية».

نشرت مسوَّدة توجيهات أكتفل عام ١٩٨٢ باسم (الإصدار المؤقت) وشملت

الاستماع والكلام والقراءة والكتابة والثقافة وكانت مقسّمة إلى أربعة مستويات رئيسة، ومعدَّلة عن إطار المائدة المستديرة بين الوكالات المعروف باسم ILR وخضعت لملاحظات ميدانية من المعنيين بتدريس اللُّغات الأجنبية ثُم صدرت بشكلٍ رسميّ لأول مرة عام ١٩٨٦م بدون القِسم الخاص بالثقافة وبخمسة مستويات رئيسة لمهارتي (الاستماع والقراءة) وأربعة مستويات لمهارتي (الكلام والكتابة)، وحملت اسم (توجيهات أكتفل)، وعلى الرغم من أنها قدمت بعدِّها إصدارًا شاملًا لكل اللُّغات الأجنبية إلا أن المجلس الأمريكي لتعليم اللُّغات الأجنبية أصدر عدة نسخ خاصة بلغات أخرى منها العربية. ثمّ طورت التوجيهات الخاصة بالكلام (١٩٩٩) والكتابة (٢٠٠١)، ثم صدر إصدار ٢٠١٢ والذي أضاف للكلام والكتابة مستوى خامس (المتميز). (الحدقي، ٢٠٢١؛ Actfl, 2012a)

صدرت من الإطار الأوربي مسودتان عام ١٩٩٦ أعقب ذلك تجريبه على نطاق واسع إلى أن خرج الإصدار النهائي الأول عام ٢٠٠١ بالإنجليزية والفرنسية، ووصلت ترجماته إلى أربعين لغة (Little, 2006; Piccardo 2020).ومع الانتشار الواسع للإصدار الأول وما أعقبه من تجريب على نطاق دولي كبير «اكتسب نوعًا من الهالة التي أدت حتمًا إلى ردود فعل معاكسة ولكن بنفس الخطورة»، حيث عده البعض وثيقة مرجعية «معقدة وغنية»(Piccardo 2020,1) . بحسب بيكاردو المؤلفة المشاركة في المجلد المصاحب.

ومع هذا الانتشار والنقد في الوقت نفسه كانت هناك حاجة لتطويره وإصدار نسخة جديدة بعنوان «المجلد المصاحب» صدرت مسودتها إلكترونيًّا عام ٢٠١٧ ثم صدرت عام ٢٠١٨ في شكل إلكتروني مؤقت بعنوان «الإطار الأوربي المرجعي المشترك للغات المجلد المصاحب مع واصفات جديدة»، ثم صدر الإصدار الثاني بشكل نهائي عام ٢٠٢٠م بعنوان «الإطار الأوربي المرجعي المشترك للُّغات المجلد المصاحب» وقد جاء أكثر سهولة ووضوحًا مُتلافيًا الكثير من الملاحظات

الـتي ذُكـرت عـبر العقديـن الماضيـين وشـاملًا لقسـم مخصّـص للغـة الإشـارة الـتي يسـتعملها الصـم والبكـم، يضـاف إلى ذلـك إصـدار كتابـين يتضمنـان عينـات ممثلـة مـن الإطـار الأوربي موجهـة للصغـار: الأول لعمـر (٧–١٠) والثـاني لعمـر (١١–١٥). (Goodier & Szabo, 2018)

ومـن ثـم فـإن كلا الإطاريـن أكتفـل والمرجعـي الأوربي لهمـا إصـدارات تجريبيـة، وصـدر كلٌّ منهمـا في إصداريـن شـاملين توجيهـات المجلـس الأمريكـي لتعليـم اللُّغـات الأجنبيـة (١٩٨٦–٢٠١٢) والإطـار الأوربي المرجعـي المشـترك للغـات (٢٠١١–٢٠٢٠).

٥-٢ التقنين:

لا يُظهـر موقـع أكتفـل كيفيـة تقنـين التوجيهـات في أي إصـدار مـن إصدارتـه. بينمـا موقـع الإطـار الأوربي يوضّـح إجـراءات التقنين.

وبينمـا لا يوضّـح موقـع أكتفـل الأُسـس الـتي اُختـيرت الواصفـات في ضوئهـا وكيفيـة تدرّجهـا. «تسـتند واصفـات الإطـار المرجعـي والمجلـد المصاحـب بشـكل أسـاسي إلى أوصـاف الخـبراء. كمـا خضعـت الواصفـات للدراسـة التجريبيـة حيـث طُلـب (مـن المعلمـين والمتخصصـين في حالـة المجلـد المصاحـب) تقديـر مسـتويات صعوبـة الواصفـات، وأُخضِعـت هـذه التقديـرات للتحليـلات الإحصائيـة، ممـا أدّى إلى مسـتويات إتقـان تصاعديـة سـتة». (Harsch & Malone, 2020, 37)

ومن الدراسات التي تناولت بناء وتقنين الإطار الأوربي بإصداريه:

North (2000, 2007); North & Piccardo (2016, 2019); North & Docherty (2016); Council of Europe, (2020)

وليست كل الواصفات المنسوبة للإطار الأوربي مُقنّنة، فالواصفات المنشورة في كتاب الإطار الأوربي مرَّت بعملية التأكُّد من صدقها وخضعت لاستشارات المعنيين بتعليم اللغة. وهناك واصفات إضافية كُتبت عن طريق المؤسسات المهتمة بالعمل في ضوء الإطار الأوربي وهذه لا تقدم أحيانًا ما يثبت تقنين واصفاتها. Council of Europe, (n.d,b)

ونخلص ممَّا سبق أنه من السهل العثور على ما يتعلق بتقنين الإطار الأوربي المرجعي المشترك للغات، بينما هذا ليس مُتاحًا في حالة توجيهات المجلس الأمريكي لتعليم اللُّغات الأجنبية.

٦-٢ سلم المستويات وأقسامه.

كما سبق وذكر الباحث فإن هذه الأُطر تتكون من بُعْدين رئيسين: البعد الرأسي هو سلم المستويات، والبعد الأفقي ما يستطيع المتعلم فعله باللُّغة.

من حيث البعد الرأسي يتكون سلم المستويات في توجيهات أكتفل من خمسة مستويات رئيسة هي (المبتدئ، والمتوسط، والمتقدم، والمتفوق، والمتميز) ويشمل كلّ مستوى من المستويات الثلاثة الأولى ثلاثة مستويات فرعية هي (الأدنى، والأوسط والأعلى)، وتُرَتَّب المستويات كلُّها من الأعلى للأدنى.

ومن حيث البعد الأفقي فينقسم إلى المهارات الأربعة (الاستماع والكلام والقراءة والكتابة).

أمّا الإطار الأوربي فينقسم إلى ثلاثة مستويات رئيسة (المبتدئ والمتوسط والمتقدم) وفي المبتدئ هناك مستويان فرعيان (١،٢)، وقد زِيد في المجلد المصاحب المستوى (التمهيدي أو ما قبل المبتدئ) ويُعَدّ جزءًا من المبتدئ الأدنى A١

كما أن هناك مستويات (+) وهي تعبر عن مستوى كفاءة أعلى من المستوى العادي لكنه لا يدرك أداء المستوى الأعلى، فمثلا مستوى (ب٢+) يعبر عن مستوى أداء أعلى من مستوى (ب٢) لكنه لا يصل لمستوى (ج١).

أما البعد الأفقي للإطار الأوربي ثلاثة أقسام رئيسة: الكفايات والأنشطة والإستراتيجيات، وتشمل ما يلي:

الكفايات اللُّغوية التواصلية: اللُّغوية، اللُّغوية الاجتماعية، التداولية

الأنشطة اللُّغوية التواصلية: الاستقبال (المقروء والمسموع)، الإنتاج (الكلام والكتابة)، التفاعل (الشفهي، والتحريري)، الوساطة.

الإستراتيجيات اللُّغوية التواصلية: الاستقبال، الإنتاج، التفاعل، الوساطة.
(Council of Europe, 2020)

نخلص ممَّا سبق إلى أنّ سلم المستويات في توجيهات المجلس الأمريكي يتكون من خمسة مستويات رئيسة مقسّمة حسب المهارات، والإطار الأوربي يتضمن ثلاثة مستويات رئيسة مقسّمة حسب الأنشطة اللُّغوية (المهارات) والكفايات والإستراتيجيات.

٢-٧ الواصفات:

تصنف واصفات توجيهات أكتفل حسب المهارات الأربعة المعروفة كما ذُكِر سابقًا، وتشمل هذه الواصفات في داخلها الجوانب المهارية والكفايات اللُّغوية واللُّغوية الاجتماعية، كما أنها توضّح جوانب القصور لدى متعلم اللُّغة في كلّ مستوى.

ومن أمثلة الواصفات (الكلام في المستوى المتقدم الأعلى):

«يتمكن الناطق في المستوى الثانوي المتقدم الأعلى من أداء كافة المهام اللُّغوية الخاصة بالمستوى المتقدم بسهولة وثقة وكفاية. إنه قادر على الشرح التفصيلي والسرد المطوّل في جميع الأزمنة بدقة لغوية. بالإضافة إلى ذلك فالمتقدم الأعلى يستطيع التعامل مع مهام لغوية تخصُّ المستوى المتميّز، إلا أنه غير قادر على المثابرة على هذا المستوى من الأداء بثبات في عدد من المواضيع المختلفة. فبإمكانه تقديم دليل متماسك لتأييد آرائه وقد يأتي بافتراضات، لكن أنماطًا من الأخطاء تبدأ بالظهور. يستطيع مناقشة موضوعات في المستوى المجرّد خصوصًا في مجال اهتماماته وتخصصه وخبرته، إلا أنه عموما أكثر ارتياحًا في مناقشة الموضوعات بالمحسوس.

يُبدي المتقدّم الأعلى مقدرة متطوّرة على تعويض النقص في تمكّنه من بعض التراكيب أو المفردات باستخدامه أساليب التواصل بثقة كإعادة صياغة أفكارِه أو الدوران حول المعنى أو الإتيان بالأمثلة، ويستخدم المفردات بدقة، ويستخدم التنغيم لإثراء المعنى ويبدي عادة طلاقة عالية ويسرًا في الكلام. إلا أنه حين يطلب منه تأدية المهام المعقدة المرتبطة بالمستوى المتميّز في عدد من الموضوعات قد تتعطّل لغته أحيانًا أو تكون قاصرة، أو قد يتجنّب المهمة كُلّيًّا باللجوء إلى التبسيط وذلك باستخدام الوصف أو السرد بدلًا من الجدل أو الافتراض». (Actfl, 2012a)

سنلاحظ في المثال السابق أنَّ الواصفات تتضمن المهارات أو الوظائف اللُّغوية «قادر على الشرح التفصيلي والسرد المطوّل-وقد يأتي بافتراضات-يستطيع مناقشة موضوعات»، وتتضمن الواصفاتُ الكفايةَ القواعدية فتقول «في جميع الأزمنة» وتحدد مستوى الجودة اللُّغوية «بدقة لغوية-بإمكانه تقديم دليل متماسك -ويبدي عادة طلاقة عالية ويسرًا في الكلام». كما توضّح الواصفات السقف الذي يمكن أن يصل إليه المتكلم «يستطيع التعامل مع مهام تخصّ المستوى

المتميز». وتحدّد الواصفاتُ نوعيةَ الخطاب عند المتكلم فتذكر «يستطيع مناقشة موضوعات في المستوى المجرّد خصوصا في مجال اهتماماته وتخصّصه وخبرته». كما توضّح الواصفات الكفاية المعجمية عند المتكلم «ويستخدم المفردات بدقة».

بالإضافة لما سبقَ توضّح الواصفات الإستراتيجيات التي يتبعها المتكلم «إعادة صياغة أفكاره أو الدوران حول المعنى أو الإتيان بالأمثلة -قد يتجنّب المهمة كُلّيًا باللجوء إلى التبسيط».

كما تحدّد الواصفات جوانب القصور لدى المتكلم في هذا المستوى «إلا أنه غير قادر على المثابرة-لكن أنماطًا من الأخطاء تبدأ بالظهور-قد تتعطل لغته أحيانًا أو تكون قاصرة».

بينما نجد الواصفات في الإطار الأوربي -كما ذكرنا سابقًا-موزّعة حسب الأنشطة اللُّغوية التواصلية والإستراتيجيات اللُّغوية التواصلية والكفايات اللُّغوية التواصلية، وكلّ قسم من هذه الأقسام الثلاثة له تفريعاته حيث تنقسم الأقسام الثلاثة الرئيسة هذه إلى ١٣ عشر قسمًا فرعيًّا، وهذه الأقسام الفرعية تبدأ بجدول واصفات عامّة ثم جداول تعبر عن أقسام أدقّ.

ومن ثم فواصفات الإطار المرجعي متشعبة ومتفرعة وتركّز على جانب محدّد مثلا: (الحديث للجمهور) (كتابة المراسلات) (الكفاءة المعجمية) (إستراتيجيات الاستقبال) وهكذا.

ومن أمثلة الواصفات (الكلام في المستوى المتوسط الأعلى):

"ب ٢+

يستطيع أن يقدّم وصفًا وعروضًا واضحة، ومطوّرة بشكل منهجي، مع الإبراز المناسب للنقاط المهمة، وللتفاصيل الداعمة ذات الصلة.

ب ٢

يستطيع أن يقدم أوصافًا واضحة ومفصّلة وعروضًا تقديمية في مجموعة كبيرة من الموضوعات المتعلقة بمجال اهتمامه؛ موسعًا للأفكار بنقاطها الفرعية وداعمًا لها بالأمثلة ذات الصلة».

(مجلس أوربا، ٢٠٢٠، ٨٣)

سنلاحظ في هذا المثال أنه يولي اهتمامًا بالنشاط التواصلي (يقدم وصفًا وعروضًا واضحة) وسنلاحظ الفرق بين المستويين حيث نجد المستوى + يتضمن جودة في العرض (بشكل منهجي، إبراز النقاط المهمة)، (التفاصيل الداعمة ذات الصلة) بينما لن نجد هذه الإمكانات في المستوى الأقل (ب٢).

في الوقت نفسه لن نجد وصفًا للكفاية القواعدية لدى الشخص ضمن واصفات الكلام بل سنحتاج أن نذهب لمقياس آخر يتضمن الكفاية القواعدية في هذا المستوى.

مثال آخر: الدّقة النَّحوية

"ب ٢+

التحكُّم النحْوي جيّد، وتحدث زلات أو أخطاء غير منهجية من حين لآخر، وقد تكون هناك عِلَل يسيرة في بنية الجملة، ولكنها نادرة ويمكن تصحيحها غالبًا بأثر رجعي.

ب ٢

يُظهر درجة عالية نسبيًّا من التحكم النحوي، ولا يرتكب أخطاء تؤدي إلى سوء الفهم.

لديه تمكّن جيّد من تراكيب اللغة السهلة، وبعض الأشكال النحوية المعقدة، مع أنه يميل إلى أن يستخدم التراكيب المعقدة بصورة صارمة بشيء من عدم الدقة». (السابق، ١٦١)

سنلاحظ في المثال السابق في هذه الواصفات تركيزًا على مدى الدقة النحْوية (التحكم النحْوي جيد، يظهر درجة عالية نسبيًّا من التحكم النحْوي ولا يرتكب أخطاء تؤدي إلى سوء الفهم). كما تُظْهِرُ الواصفات القصور النحْوي لدى الشخص (تحدث زلات، وقد تكون هناك علل يسيرة في بنية الجملة).

أما إن أردنا أن نعرف الإستراتيجيات التي يستخدمها الشخص في هذا المستوى فنحتاج أن نذهب لمقياس آخر خاص بالإستراتيجيات

مثال: إستراتيجية التخطيط

ب ٢+

«يستطيع عند الاستعداد لموقف يحتمل أن يكون معقدًا أو محرجًا، أن يخطط لما يقوله في حالة حدوث ردود فعل مختلفة، ويتأمل في التعبير الذي سيكون ملائمًا»

ب٢

يستطيع أن يخطّط لما ينبغي قوله، والوسائل لقول ذلك، مراعيًا تأثير ذلك على المتلقين.

يستطيع أن يتدرّب على توليفات وتعبيرات جديدة، ويجرب إلقاءها مما يستدعي تقديم ملاحظات». (السابق، ٩١)

تتناول هذه الواصفات إستراتيجية التخطيط في هذا المستوى (يخطّط لما يقوله، يتأمّل التعبير الملائم، يخطّط لما ينبغي قوله)

كما أنَّ الواصفات المذكورة في الإطار الأوربي المرجعي ليست كلّها لوظيفة واحدة؛ فهناك واصفات للتقييم الذاتي كما يظهر في ملحق (٢) بالمجلد المصاحب وهي موجهة للمتعلمين لتقييم أنفسهم، وواصفات موجهة للمُقَيِّمِينَ كما يظهر في ملحق (٣) الخاصّ بالميزات النوعية للُّغة المنطوقة وملحق (٤) الخاص

بالتقييم الكتابي، وواصفات موجهة لعملية التدريس وتصميم المناهج وبناء المواد التعليمية وهي الواصفات في المقاييس الأخرى في الإطار.

نخلص من هذه النقطة أنَّ الواصفات في توجيهات أكتفل تُقَدَّم في منظومة واحدة مقسمة حسب المهارات، بينما في الإطار الأوربي المرجعي تتوزع تلك الواصفات حسب الأنشطة التواصلية والإستراتيجيات والكفايات اللُّغوية التواصلية، بما يجعل القارئ يحتاج للتعامل مع عدد كبير من المقاييس في كل مستوى.

٨-٢ الدعم الفني/الخدمات:

يوفّر موقع المجلس الأمريكي لتعليم اللُّغات الأجنبية Actfl توجيهاته ٢٠١٢ مجانًا بأربع عشرة لغةً مدعومة بنماذج لمقابلات حقيقية تعبر عن مستوى الأداء في المستويات المختلفة، بالإضافة للإصدارات السابقة للتوجيهات (١٩٨٦-١٩٩٩-٢٠٠١)، وكذلك يضافُ عدد من الخدمات والموادّ بعضها مدفوع وبعضها مجاني تشمل تأهيل المعلمين. فضلًا عن تدريب المقَيِّمينَ اللغويين على إجراء الاختبارات المختلفة واعتمادهم هذا بالتعاون مع مؤسسة الاختبارات الدولية LTI التي أعطاها وكالة حصرية لإجراء الاختبارات وإصدار شهادات دولية في الكفاءة اللُّغوية طبقا لمعايير توجيهات أكتفل في ١٢٠ لغة. (Language Testing International, n.d)

ويوفّر موقع أكتفل قوائم (يستطيع أن يفعل). وينشر دورية علمية بعنوان حوليات اللغة الأجنبية Foreign Language Annals.

كما أضافت توجيهات أكتفل قسمًا إضافيًّا خاصًا يراعي خصوصية اللغة العربية وثقافتها. (Actfl, 2012b)، وهو ميزة ليست موجودة في الأُطر الأخرى في حدود علم الباحث.

على الجانب الآخر يوفّر الإطار الأوربي المرجعي المشترك للغات كتابه بإصداراته (٢٠٠١-٢٠١٨-٢٠٢٠) باللُّغات التي ترجم إليها والتي تُحَدَّثُ تِبَاعًا مجانًا، كما يوفر النُّسخ الخاصة بالمتعلمين الصغار (٧-١٠، ١١-١٥)، بالإضافة لمصادر مختلفة من إعداد مجلس أوربا ومؤسّسات تعليم اللُّغات في مجلس أوربا المهتمة بتعليم اللغة مثل المركز الأوروبي للغات الحديثة (ECML)، ورابطة مختبري اللغة في أوروبا (ALTE)، والمجلس الأوربي للغات(ELC) ، والرابطة الأوربية للتقييم والاختبارات اللُّغوية (EALTA) والرابطة الأوربية لجودة الخدمات اللُّغوية «Eaquals» التي تتولى اعتماد مؤسّسات التعليم اللّغوي. لذلك سنجد مواد تتعلق بتدريب المعنيين بالتعليم على استعمال الإطار الأوربي، وعدة أدلة للاختبارات، والدراسات المتعلقة بتقنين الإطار والجودة في التعليم اللُّغوي، كما أن هناك دورية علمية تعنى بالبحوث المتعلقة بالإطار الأوربي المرجعي بعنوان دورية سيفر CEFR Journal. (انظر جدول ١)

حتى الآن لا توجد اختبارات كفاءة لغوية طبقًا لمعايير الإطار الأوربي المرجعي معتمدة من مجلس أوربا - الجهة المالكة للإطار الأوربي-ولكن من يرغب في تقييم بناء على معايير الإطار الأوربي يعتمد على مؤسسة الاختبارات الدولية LTI التي تُقَيِّم في ضوء معادلة مستويات الإطار الأوربي بتوجيهات أكتفل؛ أي أنّ المتحَن يُمتحن وَفق مستويات أكتفل ثم تصدر النتيجة بعد معادلتها بمستويات الإطار الأوربي. كما أن هناك عددًا من المؤسسات حول العالم تنشر اختبارات وفق معايير الإطار الأوربي لكنّها ليست معتمدة من مجلس أوربا أيضا.

جدول (١) مقارنة بين الدعم الفني والخدمات التي يقدمها المجلس الأمريكي لتعليم اللُّغات الأجنبية والإطار الأوربي المرجعي المشترك للغات

	ACTFL	CEFR
إتاحة الإطار عبر الإنترنت	متاح مجانًا	متاح مجانًا
الترجمة للعربية	متاحة مجانًا	متاحة مجانًا
مادة خاصة باللغة العربية ضمن الواصفات	متاحة مجانًا	لا يوجد
نماذج عملية حقيقية من الأداء اللغوي لطلاب عبر المستويات	متاحة مجانًا	لا يوجد
الإصدارات القديمة	متاحة مجانًا	متاحة مجانًا
مصادر مجانية لتدريب المعلمين	--	متاح مجانًا
تدريب المعلمين	عبر ورشات مدفوعة	عبر مؤسسات أخرى
مصادر تصميم المناهج في ضوء الإطار	متاحة في كتب أخرى مدفوعة	متاح مجانًا + مدفوعة
مبادئ التدريس في ضوء الإطار	متاح مجانًا	متاح مجانًا
مصادر إعداد الاختبارات في ضوء الإطار	عبر ورشات مدفوعة	متاح مجانًا ب ٣ كتب
التدريب على إعداد الاختبارات	عبر ورشات مدفوعة	خدمة مدفوعة عبر طرف آخر
اعتماد المختبرين	متاح "خدمة مدفوعة"	خدمة مدفوعة عبر طرف آخر
اعتماد المؤسسات	--	متاح عبر طرف آخر
الاختبارات المعيارية المعتمدة	متاحة عبر طرف آخر "خدمة مدفوعة"	غير متاح

	ACTFL	CEFR
اعتماد الاختبارات للغير	غيرمتاح	غيرمتاح
الدوريات العلمية	Foreign Language Annals	CEFR Journal
المؤسّسات الشريكة	عشـرات مـن المؤسسـات أعضـاء في الاتحـاد	مؤسسـات تعليـم اللُّغـات في دول مجلـس أوربـا

٢-٩ الأثر

بالإضافـة لدورهـا الأسـاسي في التقييـم اللغـوي أثـرت توجيهـات أكتفـل في العمليـة التعليميـة بشـكل عـام، وظهـر ذلـك في بنـاء واصفـات أداء التلاميـذ في مراحـل التعليـم قبـل الجامعـيّ مـن ريـاض الأطفـال إلى الصـف الثـاني عشـر، كمـا تسـتخدم التوجيهـات بجانـب المعايـير الوطنيـة لتعلـم اللُّغـات الأجنبيـة -بإصداراتهـا المتعـددة- لوصـف مـدى تلبيـة الطـلاب لمعايـير محتـوى المقـررات الدراسـية اللغويـة. (Actfl, 2012a).

كمـا أن العديـد مـن الأُطـر المرجعيـة الدوليـة مثـل الإطـار الأوربي تعتمـد علـى توجيهـات أكتفـل في الاختبـارات اللُّغويـة - كمـا ذُكـر سـابقًا-واعتمد إطـار حلـف شـمال الأطلسـي علـى أكتفـل في إعـداد اختباراتـه وعلـى مؤسّسـة الاختبـارات الدوليـة في تنفيذهـا. (الحـدقي، ٢٠٢١).

يُعَـدّ الإطـار الأوربي المرجعـي دسـتور تعليـم اللُّغـات في دول مجلـس أوربـا الـستة والأربعـين (Council of Europe, n.d,a)، كمـا يسـتخدم علـى نطـاق واسـع حـول العالم وهـذا يظهـر في تعـدد ترجماتـه الـتي تجـاوزت ٤٠ لغـة. (Figueras, 2012)

٢-١٠ النقد

بداية لا يوفر موقع أكتفل معلومات عن كيفية تطوير توجيهاته أو إجراءات تقنينها، وعند البحث في محرك البحث العلمي (باحث جوجل) لا نصل لنتيجة مرضية؛ حيث تتركز أغلب البحوث لتقييم كفاءة اختباراته.

أُخِذ على توجيهات أكتفل عدم استنادها لنظرية ما أو توضيح النظريات التي اعتمدت عليها في بنائها لإيضاح تصورها لمكونات الكفاءة اللُّغوية وسلم اكتساب الإنسان للغة. (Chalhoub-Deville, 1997)، وعلى الرغم من قدم هذا النقد إلا أن موقع توجيهات أكتفل في إصدارها الأحدث ٢٠١٢ لا يقدم توضيحات أو ردًا عليه. (الحدقي، ٢٠٢١)

كما أُخِذ على أكتفل وضع الناطق الأصلي جيد التعليم معيارًا مثاليًّا للأداء اللغوي؛ حيث يُعَدّ هذا المعيار غير منطقي فضلًا عن مطاطيته، فالناطق الأصلي لديه كمٌّ هائل من المُدخلات اللُّغوية والثقافية التي يكتسبها منذ ميلاده ومن غير المنطقي مقارنة متعلِّم اللغة الأجنبية به، كما أنَّ وصف «جيد التعليم» غير محدّد فمن هو جيد التعليم؟ ومن هو الذي لم يتعلم جيدًا؟

نلاحظ أيضا أنَّ توجيهات أكتفل لا تتعرض لموضوعات تهمّ متعلمي اللغة مثل الترجمة والتفاعل عبر الإنترنت، وكأنها لا تُجَارِي سرعة نمو الأنشطة التواصلية في هذا العصر شديد التغير.

أما فيما يخص الإطار الأوربي المرجعي فالبعض يرى أنَّ الأساس النظري لمعيار CEFR أحد نقاط ضعفه (على سبيل المثال Hulstijn ، 2007) وقد ردَّ على هذا النقد المؤلفُ الرئيسُ للإطار الأوربي نورث في مقال خاصّ بأن الإطار يستند لنماذج معروفة شهيرة وهي نموذج القدرة اللُّغوية التواصلية (١٩٩٦) ونموذج الكفاية التواصلية (١٩٩٥). (North, 2020) كما ذكرنا من قبل.

تذكر (Harsch, 2014, 156) أنَّ واصفات الإصدار الأول للإطار الأوربي تستند إلى «الإجماع المهني بدلًا من الأدلة التجريبية». بينما يرى المؤلِّفُ الرئيسُ للإطار

الأوربي (٢٠٢٠) North أن إشراك الممارسين في تطوير مثلِ هذا الإطار أكثر أهمية من تأسيسه على نظريات اكتساب اللُغة. وتعلق Harsch&Malone على هذه النقطة بأن هرمية المستويات حُدّدت بناء على تقدير الخبراء والتحليلات الكَمّية وليست وصفًا مصدقًا لعملية اكتساب اللغة الثانية، كما أنه لا يصف عملية تطور الكفاءة اللُغوية لدى الأفراد بدقة. (٢٠٢٠)

وفي تقرير غير منشور رصد المؤلفون خَللًا في صياغة بعض الواصفات في إصدار ٢٠٠١ حيث استخدمت ألفاظ غير دقيقة في الواصفات مثل (بطيء، سريع، مألوف، ملموس) كما استخدمت ألفاظ في سياقات دون إيضاح هل هي مترادفات أم هي مقصودة لذاتها مثل (يفهم، يأخذ، يحدد،... إلخ).

Alderson, J. C., Figueras, N., & Kuijper, H. (2004)

ومن خلال تجربة الباحث في المنتديات والأنشطة العلمية كانت هناك شكوى متكررة من الباحثين والمعلمين من صعوبة التعامل مع كتاب الإطار الأوربي (٢٧٨صفحة) مقارنة بعدد محدود من الصفحات تحتوي توجيهات المجلس الأمريكي لتعليم اللُغات الأجنبية (٢٤ صفحة). وهذا يدفع جمهور الإطار الأوربي للجوء إلى أقسام بعينها وتجاهل الأقسام الأخرى ففي عام ٢٠٠٥ أجرى قسم سياسة اللُغات في مجلس أوروبا استبيانًا لمعرفة مدى استخدام الإطار الأوربي في مجلس أوربا جاءه ١١١ ردًا من ٣٩ دولة أكدت على أن المعرفة بالإطار الأوربي واستخدامه تقتصر على فئة محدودة من المتخصصين. وأن الأجزاء الأكثر استخدامًا منه هي المقياس العام الذي يتضمن المستويات الستة وشبكة التقييم الذاتي. (Little, 2006). وربما تكون هذه النتيجة تعبّر عن فترة قريبة زمنيًا من بداية نشر الإطار الأوربي (٢٠٠١) وقد يكون الأمر تغيرَ بعد مرور قرابة ١٨ عامًا في ظلّ الدعم الفني الكبير الذي تقوم به المؤسسات المعنية بتعليم اللُغات وفق الإطار إلا أنه يظل كبير الحجم مقارنة بتوجيهات أكتفل.

٣ - نتائج البحث ومناقشتها

في ضوء ما سبق يمكن الوصول للنتائج التالية:

٣-١ جاءت نشأة توجيهات المجلس الأمريكي لتعليم اللُّغات الأجنبية مبكرة خمسة عشر عامًا عن الإطار الأوربي المرجعي، بينما كانت الأولى نسخة معدلة من إطار المائدة المستديرة بين الوكالات ILR فإن الإطار الأوربي المرجعي بني على تجارب مجلس أوربا السابقة عليه مستفيدًا بالكثير من تجارب أُطر أخرى وصلت لثلاثين إطارًا، وقد يقول قائل ما دور الزمن في هذه المقارنة؟ إن دور الزمن مهم؛ حيث كلما قدم الإطار كانت له فرصة التجريب في سياقات كثيرة ومتعددة فضلًا عن انتشاره، وفي الوقت نفسه كُلّما كان الإطار حديثًا استوعب مستجدات الحياة وما وصل إليه العلم في قضية تعلُّم واكتساب اللغة، وأكبر دليل على دور الزمن أننا نجد أن اختبارات أكتفل للكفاءة اللُّغوية تنعقد لـ ١٢٠ لغة كما سبق وذكرنا نقلًا عن LTI، وهذا الإنجاز لم نجد ما يدل على تحقيق الإطار الأوربي له، كما أنّ حداثة الإطار الأوربي سمحت له أن يواكب المستجدات التقنية فخصّص قسما للتفاعل عبر الإنترنت وهذا أمر ليس متوفرًا في توجيهات أكتفل التي صدرت منذُ عشرة أعوام ولم تكن فيها وسائل التواصل الاجتماعي بهذه السطوة الحالية.

اتفقت توجيهات المجلس الأمريكي والإطار الأوربي أنها موجهة للمتعلمين الكبار، وقد ساعدت توجيهات أكتفل في تطوير إصدار خاص بعنوان توجيهات أكتفل لأداء المتعلمين في التعليم قبل الجامعي ACTFL Performance Guidelines for K-12 Learners. (Actfl, 2012a). كما صدرت واصفات خاصة من الإطار الأوربي لعمر (٧-١٠) و(١١-١٥)، وهذه السلسلة من الموادّ التي تغطي جميع مراحل التعليم شديدة الأهمية للتعليم اللُّغوي حيث تتحول عملية تعليم اللغة عبر المراحل التعليمية إلى منظومة مترابطة وَفق فلسفة إطار محدد.

٣-٢ الهدف الرئيس لتوجيهات المجلس الأمريكي هو التقييم اللُّغوي، بينما الإطار الأوربي أكثر اتساعًا للتعليم والتعلّم والتقييم، ويرى الباحث أنه من الطبيعي كما تقدم الأُطرُ واصفاتٍ لكافة الأعمار عليها أن تقدم -أيضا- منظومة تغطي العملية التعليمية حتّى لا نجد أن المناهج تغرد في طريق يختلف عما تقيسه الاختبارات أو العكس.

جدول (٢) موجز نتائج البحث

	أكتفل	الإطار الأوربي
النشأة	١٩٨٦	٢٠٠١
الجمهور	الكبار (لكن هناك توجيهات أخرى للتعليم قبل الجامعي)	الكبار (لكن هناك إصدارين للصغار)
الأهداف	التقييم	التعلُّم، التعليم، التقييم
النظرية	لا توجد نظرية محدّدة	نظرية القدرة اللُّغوية التواصلية والنموذج التربوي للكفاية التواصلية.
الإصدارات	١٩٨٦ – ٢٠١٢	٢٠٠١ – ٢٠٢٠
التطوير	جزئيًّا خلال ١٣ – ١٥ عاما كُلّيًّا خلال ٣٠ عاما	مرة كُلّيًّا خلال ١٧ عاما
التقنين	غير مذكور في الموقع	مذكور في الإطار ومصادر أخرى
سلم المستويات	٥ رئيسة، ١١ مستوى	٣ رئيسة، ٦ فرعية
الواصفات	مقسَّمة حسب المهارات وتتضمن في داخلها المهارات والكفايات والإستراتيجيات	مقسَّمة حسب الأنشطة اللُّغوية والكفايات والإستراتيجيات وكل مجموعة واصفات تتناول جانبًا واحدًا من تلك الأقسام
الدعم الفني والخدمات	بعضه مجانيّ وأغلبه مدفوع	المصادر المجانية كثيرة، وهناك مصادر مدفوعة
الأثر	موجود	موجود
النقد	عدم استناده لنظرية، عدُّ الناطق الأصلي معيارًا، عدم وضوح إجراءات التقنين	ضعف أساس النظرية، تشعبه وكبر حجمه، لا يوجد له اختبارات معتمدة

٣-٣ لم تستند توجيهات المجلس الأمريكي لتعليم اللُّغات الأجنبية إلى أيّ نظرية ونصّت على ذلك في مقدمة التوجيهات، بينما استعان الإطار بنموذجي القدرة اللُّغوية التواصلية (١٩٩٦) والنموذج التربوي للكفاية التواصلية (١٩٩٥).

رأت توجيهات المجلس الأمريكي الناطقَ الأصلي جيد التعليم هو المعيار المثالي للأداء اللُّغوي والذي يُنسب إليه أداء المتعلمين، بينما أعدَّ الإطار الأوربي المتعلمَ وسيطًا اجتماعيًّا يقاس حسب قيامه بالمهام التواصلية المحدّدة في كلّ مستوى.

٣-٤ صدر الإطاران في شكلٍ تجريبي أولًا ثم أعقب ذلك صدور نسخ نهائية، وخرج الإصدار الأول لتوجيهات المجلس الأمريكي عام ١٩٨٦ ثم الإصدار الثاني عام ٢٠١٢، وكان هناك تطوير جزئيّ لمهارات الإنتاج عامي ١٩٩٩ و٢٠٠١م، والحال مع الإطار الأوربي لا يختلف كثيرًا حيث صدرت نسخته الأولى عام ٢٠٠١ ونسخته الثانية عام ٢٠١٨م، وعمليات التطوير هذه هي التي تضمن لهذه الأُطر حيويتها ومواكبتها لمستجدات العصر العلمية والعلمية.

٣-٥ لا يذكر موقع المجلس الأمريكي طريقة تقنين توجيهات الكفاءة اللُّغوية، وقد عاب الباحثون هذا الأمر بينما يذكُر عدد من الدراسات كيفية تقنين الإطار الأوربي المرجعي. ويرى الباحث أن قضية تقنين الأُطر أمر شديد الأهمية لما يترتب عليها؛ حيث يُعتمَد عليها في عملية التعليم اللُّغوي على نطاق واسع الآن والتعامل معها دون التأكُّد من معياريتها قد يطعن في كفاءة نواتجها.

٣-٦ عدد مستويات توجيهات أكتفل ٥ رئيسة تشمل (١١ مستوى)، وهذه المستويات مُقسَّمة حسب المهارات الأربعة، بينما عدد مستويات الإطار الأوربي الرئيسة ثلاثة تشمل ستة مستويات فرعية، إلا أنّ هناك بعض المستويات الفرعية تنقسم إلى مستويين داخليين ويسمى الأعلى منها + -وهو أمر متعارف عليه في الأُطر المرجعية واستخدمه إطار المائدة المستديرة ILR- والمستويات في الإطار الأوربي مقسمة حسب الأنشطة اللُّغوية (الاستقبال والإنتاج والتفاعل

والوساطة) وحسب الإستراتيجيات للأقسام السابقة وحسب الكفايات (اللُّغوية، اللُّغوية الاجتماعية، التداولية).

٣-٧ هناك فَرْقٌ واضح بين حجم المادة التي يتعامل معها مستخدم توجيهات أكتفل والتي تساوي ٢٤صفحة وبين مقاييس الإطار الأوربي والتي يحتويها كتاب في ثلاثمائة صفحة. وهذا الفرق الكبير يلفت النظر؛ ففرق كبير بين شخص يحتاج أن يتعلم أو يتدرب على ٢٤صفحة وشخص آخر يحتاج أن يتعامل مع قرابة الثلاثمائة صفحة، وهذا يظهر على جميع الأصعدة الاستيعاب والتطبيق والقرارات.

إلا أننا يمكن أن ننظر للأمر من زاوية أخرى وهو أنَّ الإطار الأوربيّ مصمَّم لمجالات أكثر اتساعًا من أكتفل؛ فهو موجَّه للطالب والمعلم والمقَيِّم (تعليم تعلُّم تقييم) بينما توجيهات أكتفل مصمَّمة للتقييم فحسب وإن تمَّ الاستعانة بها في بناء توجيهات أخرى تختصّ بالتعليم والتعلُّم، ومن ثم فمن الطبيعي أن يكون الإطار الأوربي بهذا الحجم، كما أنَّ وفرة الواصفات وتخصّصها يساعد مُصمِّم المناهج والموادّ التعليمية للتركيز على جزئيات بعينها وتأليف ما يناسبها أو تقييمه، بينما واصفات أكتفل لا توفر ذلك؛ حيث تتركز حول تقييم الأداء اللُّغوي. كما أن هناك جوانب في واصفات الإطار الأوربي لم تتطرّق لها توجيهات أكتفل مثل الوساطة والتفاعل عبر الإنترنت والتوازي اللُّغوي

٣-٨ يوفر كلا الإطارين خدماتٍ ودعمًا فنيًّا في مختلف جوانب العملية التعليمية، وأهم ما يميز توجيهات أكتفل وجود قسم خاصّ باللغة العربية -وهذا غير الترجمة العربية- كما أنَّ هناك العديد من المقابلات الشخصية الشفهية المسموعة التي تقدّم أداء ممثلًا للمستويات المختلفة، وكلاهما لا يوفّره الإطار الأوربي. بينما يوفّر الإطار الأوربي موادّ مجانية من إعداد مجلس أوربا ومنها من إعداد الجهات شريكة المجلس فضلًا عن الموادّ المحلية التي يُنتجها أهل كلّ لغة. وهو ما يجعل المصادر المجانية مُتاحة لكل من أرد العمل أو البحث في ضوء الإطار الأوربي.

كما تتميز توجيهات أكتفل بوجود اختبار مقنن عالميّ ممثِّل لها في مختلف لغات العالم؛ بينما لا يتوفر هذا للإطار الأوربي حيث لا يوجد اختبار معتمد من مجلس أوربا، ولكن هناك مؤسّسات تقدم هذه الاختبارات بناء على مسئوليتها.

٣-٩ كلا الإطارين له أثر كبير في مجال تعليم اللُّغات الأجنبية، فتوجيهات أكتفل قادت التقييم الأكاديمي في الغرب لفترة ليست قصيرة ومازالت، فضلا عن اعتماد مؤسّسات دولية عليها في بناء نظمها للتقييم مثل إطار حلف شمال الأطلسي. أمّا الإطار الأوربي فصار مرجعية لدول مجلس أوربا الـ ٤٦ بالإضافة للمؤسّسات التي تبنَّت الإطار نظامًا لها في كثير من دول العالم، لدرجة أنّ دورية الإطار الأوربي تصدر عن رابطة لغوية في اليابان وهي ليست عضوًا في مجلس أوربا.

٣-١٠ وُجّه نقد لكلا الإطارين؛ وأكثر ما وجه لتوجيهات أكتفل هو عدم اعتمادها على نظرية، وعدها الناطق الأصلي معيارًا للأداء اللغوي فضلًا عن عدم وضوح طريقة تقنينها. أمّا الإطار الأوربي فقد وُجّه له نقد باعتماده على تقديرات المختصين أكثر من اعتماده على نظريات اكتساب اللُّغة، فضلًا عن كبر حجمه وتشعبه ومطاطية ألفاظه، وإن كانت هذه الملاحظات ردّ عليها نظريًّا براين نورث في مقال (North, 2020) وبعضها رد عليه مؤلفو الإطار عمليًّا بتغيير العديد من واصفات الإطار في نسخته الأخيرة «المجلد المصاحب».

اختبارات اللغة
Journal of Language Testing

٤ - الخاتمة

حاول هذا المقال المقارنة بين إطارين مهمين للكفاءة اللُّغوية -إن لم يكونا الأهمين- وهما توجيهات المجلس الأمريكي لتعليم اللُّغات الأجنبية والإطار الأوربي المرجعيّ المشترك للغات (تعلُّم، تعليم، تقييم)، بهدف الوقوف على نقاط الاتفاق والاختلاف، وقد ركّز هذا المقال على خصائص كلّ إطار وما يوفّره من دعم لمستخدميه وعرض ما وُجّه إليه من نقد في سبيل مساعدة المهتمين بالأُطر المرجعية للكفاءة اللُّغوية على اختيار المناسب منها لحاجات طلابهم ولطبيعة البيئة التعليمية التي يعملون بها.

أظهرت أهم النتائج أن الإطارين يتفقان في بنائهما على أعمال سابقة وأنهما وُجّها للجمهور الأكاديمي، وقد خضعا للتجريب قبل إصدارهما لمرتين بشكل نهائي، كما أنّ كلّ جهة من الجهات الراعية للإطارين قدمت عددًا من الخدمات المجانية والمدفوعة لمُستخدمي الإطار وإن كانت المصادر المتاحة بشكل مجاني للإطار الأوربي أكثر وفرةً، وقد يرجع ذلك لكثرة الدول التي تتبنَّى هذا الإطار. كما اتفق الإطاران في هدف هو استعمالهما في التقييم اللُّغوي. إلا أنَّ الإطار الأوربي زاد على تلك الأهداف هدفين هما التعليم والتعلُّم.

اختلفت توجيهات أكتفل عن الإطار الأوربي بعدم اعتمادها على نظرية محدَّدة في بنائها، كما أنه من غير المنشور على موقع أكتفل طريقة تقنينها. بينما قدَّم مؤلفو الإطار الأوربي النظرية التي اعتمدوا عليها في بناء مخطّطهم الوصفي كما قدموا طريقة بناء الإطار وتقنينه.

اختلفت توجيهات أكتفل عن الإطار الأوربي في حجمها، حيث جاءت بنسبة ١٠٪ تقريبا من حجم الإطار الأوربي وهو ما عَدَّه البعض ميزة تُسهّل استيعابه، بينما يمكِنُ النظر لحجم الإطار الأوربي على أنّ هذا يوفّر مخزونًا من الواصفات والأفكار التي يمكن بناء المناهج والمواد التعليمية في ضوئها.

يقترح البحث الاستفادة بهذه الميزات والنقد الموجه لهذه التجارب العالمية الكبرى عند تبنّي أحد الإطارين أو عند بناء إطار عربي منبثق من أحدهما أو حتى عند بناء إطار عربي معياري مستقل.

المراجع:

العربية:

- الحدقي. إسلام يسري. (٢٠٢١). الكفاءة اللغوية نماذجها وأطرها. أكدم
- مجلس أوروبا. (٢٠٢٠). الإطار المرجعيّ الأوروبي المشترك لتعلم اللُّغات وتعليمها وتقييمها المجلد المصاحب (عبد الناصر عثمان صبير، ترجمة؛ ط١) معهد اللغة العربية جامعة أم القرى. (٢٠٢٢)
- محمود، مصطفى عرابي. (٢٠٢٠). إطار معياري مقترح لمتعلمي اللغة العربية الناطقين بلغات أخرى. مجلة القراءة والمعرفة. م ٢٠ (الجزء الثاني مارس) ٢٣١-٢٨٣

English:

- ACTFL. (2016). Assigning CEFR ratings to ACTFL assessments. Washington, DC: American Council on the Teaching of Foreign Languages.
- Alderson, J. C., Figueras, N., & Kuijper, H. (2004). Final Report of the Dutch CEF Construct Project. Unpublished Document.
- Bachman, L. F., & Palmer, A. S. (1996). Language testing in practice: Designing and developing useful language tests (Vol. 1). Oxford University Press.
- Celce-Murcia, M., Dörnyei, Z., & Thurrell, S. (1995). Communicative competence: A pedagogically motivated model with content specifications. Issues in Applied linguistics, 6(2), 5-35.
- Chalhoub-Deville, M. (1997) Theoretical Models, Assessment Frameworks and Test Construction, Language Testing, 14,1: 3-22

- Council of Europe.(2020). Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Companion Volume, 2nd edition, Strasbourg: Language Policy Program, Council of Europe Publishing. https://rm.coe.int/common-european-framework-of-reference-for-languages-learning-teaching/16809ea0d4
- Figueras, N. (2012). The impact of the CEFR. ELT journal, 66(4), 477-485.
- Goodier, T., & Szabo, T. (2018). Collated representative simples of descriptors of language competences developed for young learners. Zurich: Council of Europe.
- Harsch, C., & Malone, M. E. (2020). Language Proficiency Frameworks and Scales. The Routledge Handbook of Second Language Acquisition and Language Testing
- Hulstijn, J. H. (2007). The shaky ground beneath the CEFR: Quantitative and qualitative dimensions of language proficiency. The Modern Language Journal, 91(4), 663-667.
- Liskin‐Gasparro, J. (1984). The ACTFL proficiency guidelines: A historical perspective. In T. Higgs (Ed.), Teaching for proficiency, the organizing principle (pp. 11–42). Lincolnwood, IL: National Textbook Co.
- Little, D. (2006). The Common European Framework of Reference for Languages: Content, purpose, origin, reception and impact. Language Teaching, 39(03), 167. doi:10.1017/s0261444806003557
- North, B. & Piccardo, E. (2016). Developing illustrative descriptors of aspects of mediation for the Common European Framework of Reference (CEFR). A Council of Europe project. Language Teaching 49(3), 1-5.
- North, B. & Piccardo,E.(2019). Developing new CEFR descriptor scales and expanding the existing ones: constructs, approaches and methodologies

- North, B. (1995). Scales of language proficiency. Melbourne Papers in Language Testing, 4(2), 60-111.
- North, B. (2000). The Development of a Common Framework Scale of Language Proficiency. New York, United States of America: Peter Lang Verlag.
- North, B. (2002). Developing descriptor scales of language proficiency for the CEF common reference levels. In J. C. Alderson (Ed.), Common European framework of reference for languages: Learning, teaching, assessment. Case studies (pp. 87–105). Council of Europe. https://rm.coe.int/168069f403
- North, B. (2007). The CEFR: Development, theoretical and practical issues. Babylonia, 1(07), 22-29.
- North, B. (2020). Trolls, unicorns and the CEFR: Precision and professionalism in criticism of the CEFR. CEFR Journal Research and Practice, 2, 8-24.
- North, B., & Docherty, C. (2016). Validating a set of CEFR illustrative descriptors for mediation. Cambridge English Press: research notes, (63), 24-33.
- North, B., & Piccardo, E. (2018). Aligning the Canadian Language Benchmarks (CLB) to the Common European Framework of References (CERF).
- North, B., Piccardo, E., Goodier, T., Fasoglio, D., Margonis-Pasinetti, R., & Rüschoff, B. (2022). Enriching 21st-Century Language Education. The CEFR Companion volume in practice.
- North, Brian. (2000). The development of a common framework scale of language proficiency. New York: Peter Lang.

- Piccardo, E. (2020). The Common European Framework of Reference (CEFR) in language education: Past, present, and future. TIRF: Language Education in Review Series, 15.
- Tschirner, E. (Ed.). (2012). Aligning frameworks of reference in language testing: The ACTFL proficiency guidelines and the common European framework of reference. Stauffenburg Verlag

Websites:

- ACTFL.(1989). ACTFL PROFICIENCY GUIDELINES version 1986. https://www.actfl.org/sites/default/files/guidelines/ACTFLProficiencyGuidelines1986.pdf
- (ACTFL). (2012a). ACTFL PROFICIENCY GUIDELINES. https://www.actfl.org/resources/actfl-proficiency-guidelines-2012
- Actfl.(2012b(. Arabic Consensus Project. https://www.actfl.org/resources/actfl-proficiency-guidelines-2012/arabic/arabic-consensus-project/speaking
- Council of Europe, (n.d,a). The Council of Europe in brief/ Our member States https://www.coe.int/en/web/about-us/our-member-states
- Council of Europe, (n.d,b). Descriptors. https://www.coe.int/en/web/common-european-framework-reference-languages/the-cefr-descriptors#
- Herzog, Martha.(n.d).An overview of the history of the ILR Language proficiency skill level descriptions and scale. HOW DID THE LANGUAGE PROFICIENCY SCALE GET STARTED?. Retrieved on 15/01/2021 from http://www.govtilr.org/Skills/IRL%20Scale%20History.htm
- Language Testing International. (n.d). About LTI/ Who we are. https://www.languagetesting.com/lti-information

الاختبارات الموضوعية الأكثر استخدامًا والمستويات والنواتج التي تستهدفها

د. محمد عبد الخالق محمد فضل[1]

مستخلص

تتناول الدراسة بالتحليل والتقييم، أكثَربنود الاختبارات الموضوعية شيوعًا، والمستويات المعرفية الأكثَراستخدامًا في تلك الاختبارات، والنواتج التعليمية التي ينبغي أن تحققها، بالإضافة إلى العيوب الكامنة فيها، وأوجه القصور التي تعتري مختلف أنواع الاختبارات الموضوعية كلًا على حدة.

خلصت الدراسة إلى أن أكثر الأنواع استخدامًا هي بنود الاختيار من متعدد، وأن هذه الاختبارات، تقتصر على قياس الأهداف المعرفية في أدنى سلم المجال المعرفي ليس إلا؛ ممثلة في الفهم والتذكر، وعلى نحو قليل مستوى التطبيق، وأقل من ذلك بكثير، مستويَيْ التحليل والتركيب، إلَّا أنها لم تقس قط المستويات المعرفية العليا الأكثر تعقيدًا التي تحتاج إلى إعمال للفكر والاستنتاج والاستنباط؛ كالتحليل والتقويم والإبداع.

تعزو الدراسة أوجه القصور والعيوب إلى عدم التزام معدي هذه الاختبارات بالمعايير العلمية التي وضعها الخبراء والتربويون في القياس التربوي والتعليمي. واتضح أن من أبرز العيوب التي استنتجتها الدراسة، أن معظم أنواع الاختبارات الموضوعية – قيد الدراسة – لا تحقق النواتج التعليمية التي يفترض أن تقيسها. وتقدم الدراسة في خاتمتها النتائجَ، وبعض التوصيات التي قد تفيد المهتمين بأمر هذا النوع من الاختبارات.

---

(١) أستاذ مساعد بجامعة الملك سعود والجامعة السعودية الإلكترونية (سابقًا)

اختبارات اللغة
Journal of Language Testing

الكلمات المفتاحية:

اختبـارات موضوعيـة، تعليـم العربيـة للناطقـين بلغـات أخـرى، تقييـم لغـوي ، سلاسـل تعليميـة، مسـتويات معرفيـة، نـواتج تعليميـة.

Abstract

Commonly used objective tests and the levels and outcomes they target.

This is an analytical and evaluative study of the most common objective test items, as well as the cognitive levels most used in those tests, and the educational outcomes that they should achieve, in addition to the inherent defects in them, and the shortcomings that characterize the various types of objective tests.

The study concluded that the most used types are multiple-choice item formats, and that these tests are limited to measuring cognitive goals at the lowest level of the cognitive domain in Bloom's Taxonomy, however, it has never measured the higher, more complex cognitive levels that need the implementation of thought, and deduction, such as analysis, evaluation, and creativity.

The study ascribes and attributes the shortcomings and defects to the non-compliance of those who prepared these tests with the scientific standards set by experts and educational and measurement scholars.

Most of the test item formats, under study, do not achieve the learning outcomes that they are supposed to measure.

In its conclusion, the study presents the results and some recommendations that may benefit those interested in these types of tests.

key words:

Objective tests , Teaching Arabic to speakers of other languages , Language assessment , Educational series , Cognitive levels, Educational outcomes.

مقدمة

حدث تحول في عمليات القياس والتقييم والتقويم وأساليبها؛ من كونها إجراءً يتم لقياس ما قد تعلمه الطلاب عقب العملية التعليمية، إلى جعلها جزءًا متكاملًا من عملية التعليم والتعلم ورافدًا لها، وتغير لدى كثير من مدرسي اللغات بالتالي، مفهومُ كثير من المصطلحات من قبيل: دافعية الطلاب والتعلم المعرفي، وصنع القرار، ونواتج التعلم. كما تغير مفهوم ما نعرفه عن أهمية عملية التقييم في التدريس الفعال، وأصبح بعض ذوي الدربة من المدرسين يقيِّمون طلابهم وَفق الأهداف التعليمية التي تتوخاها برامجهم التعليمية، وترتب على ذلك أن صارت عملية تقييم الطلاب، لا توثق ما يعرفه الطلاب وما يمكن أن يؤدوه فحسب، لكنها أصبحت تؤثر أيضًا في عملية التعلم وتوجهها. وبذلك اكتسب التقييم الذي يعزز التعلم، أهمية مماثلة - تمامًا - للتقييم الذي يوثق للتعلم.

من الملاحظ أن بعضًا ممن يُعِدّون الاختبارات الموضوعية، لا يملكون إلا حدسهم عند كتابة هذه البنود، بل إن منهم من يصمم اختباراته كما عهدها لدى من سبقه من أساتذته وغيرهم؛ لذلك تجد أن أكثر أنواع الاختبارات الموضوعية استخدامًا هي البنود التي تبدو- ظاهريًا - سهلة الإعداد، سريعة سهلة التصحيح من قبيل أسئلة الاختيار من متعدد، والصواب والخطأ... وهلم جرا. وأن هذه الأنواع من البنود والأسئلة الاختبارية تقتصر- كما وضح في هذه الدراسة ونظائرها- على قياس الأهداف المعرفية في أدنى سلم المجال المعرفي مثل: الفهم والتذكر والتطبيق، وقلما تتناول مستوى التركيب، لكنها لم تقس قط المستويات المعرفية العليا الأكثر تعقيدًا التي تحتاج إلى إعمال للفكر والاستنتاج والاستنباط؛ كالتحليل والتقويم والإبداع.

يشير طعيمة (٢٠٠٣) إلى أنه «قد أثبتت الدراسات بأن الثبات على الأساليب التقليدية في التقويم، والركون إلى الاختبارات المقالية وحدها، والعجز عن وضع متكامل

لقياس مختلف أهداف منهج اللغة العربية ... كل هذا كان كفيلًا بأن يحد من فعالية التقويم كبعد رئيس من أبعاد المنهج، وكأداة رئيسة من أدوات التطوير» (ص ٢٤).

انطلاقًا مما تقدم، قامت هذه الدراسة، بالوقوف على مواطن الخلل في الاختبارات الموضوعية التي تستخدم في مؤسسات تعليم العربية للناطقين بلغات أخرى، وألقت الضوء على أكثرها شيوعًا، وكذلك على المستويات المعرفية الأكثر استخدامًا، والنواتج التعليمية التي تحققها، بالإضافة إلى أوجه القصور الكامنة فيها، وتقديم توصيات عملية للارتقاء بمهارات من يقومون بإعدادها وتصميمها وبنائها. وأوصت بأهمية تنوع مستويات الأسئلة، وتركيزها على النواتج التعليمية لما لذلك من تأثير على عملية التعليم.

تتناول الدراسة بالتحليل والتقييم الاختبارات الموضوعية الشائعة. وأشارت إلى مزيدٍ من العناية بها حتى تعكس نواتجها الأهداف التعليمية للبرامج التعليمية التعلمية التي تتبناها؛ حيث إن استخدامها أصبح في اضطراد مستمر. وشمل البحث في ذلك: عناصر اللغة ومهاراتها في الاختبارات قيد الدراسة في المستويات: المبتدئة والمتوسطة والمتقدمة.

ولكي تحقق الدراسة ذلك، نظرت في بعض الدراسات والأدبيات التي تناولت الشيوع في الاختبارات الموضوعية، وتلك التي تناولت النواتج التعليمية والمستويات المعرفية، ثم عرضت بعض النتائج والتوصيات التي رأت أن يُستأنَس بها في تصميم الاختبار وبنائه وإعداده.

الإحساس بالمشكلة

يلمس بعض الباحثين قصورًا واضحًا في الاختبارات الموضوعية التي أصبحت تستخدم على نطاقٍ واسع في برامج تعليم العربية للناطقين بلغات أخرى. يكمن هذا القصور في: التصميم والبناء والإعداد، ويتضح بقدر كبير في نواتج التعلم.

ويعـزى ذلـك القصـور لأسـباب منهـا: صعوبـة كتابـة هـذه الاختبـارات وتصميمهـا وإعدادهـا وبنائهـا، أو قصـور في كفايـات المعلمـين الذيـن لـم ينالـوا قسـطًا مـن التدريـب فيهـا. وقـد يعـود السـبب إلى عـدم وضـوح الرؤيـة لـدى مـن يُقْـدم علـى طـرح هـذه الاختبـارات، دون ربْطِهـا بأهـداف البرامـج؛ ممـا يجعلهـا أبعـد مـن تحقيـق الأهـداف المنشـودة للبرامـج.

مشكلة الدراسة

لاحـظ الـدراس أن هنـاك عيوبًـا تصاحـب اسـتخدام الأنـواع المختلفـة للاختبـارات الموضوعيـة، وتقـف عائقًـا أمـام الاسـتخدام الأمثـل لهـا. كمـا لاحـظ أن ثمـة تضاربًـا في نتـائج المعلومـات بـين الدراسـات السـابقة أو نقصًـا في المعلومـات. ومـن قبيـل هـذا التضـارب أن هنـاك مـن الدراسـات مـا يشـير إلى أن بنـود الاختيـار مـن متعـدد هـي أكـثر الأسـئلة الموضوعيـة شـيوعًا، ومنهـا مـا يشـير إلى أن أكثرهـا شـيوعًا أسـئلة التكميـل وليسـت بنـود الاختيـار مـن متعـدد. ومنهـا مـا يذكـر أنهـا تقيـس نـواتج التعلـم بكفـاءة، ومنهـا مـا يشـير إلى أنهـا لا تصلـح لقيـاس مهـارات التفكيـر العليـا، ومنهـا مـا يناقـض ذلـك بأنـه يمكـن اسـتخدامها في قيـاس مهـارات التفكيـر العليـا علـى نحـو جيـد. ومنهـا مـا يضـع مسـتوى التذكـر في المرتبـة الأولى، ومنهـا مـا يضـع مسـتوى الفهـم في المقـام الأول.

يـرى الـدراس أنـه إذا الـتَزَم كل مـن يقـوم بمهمـة التصميـم والبنـاء والإعـداد، بربـط الاختبـار بنـواتج التعلـم، والعمـل وَفْـقَ المعايـير العلميـة، لتحققـت نتـائج القيـاس المرجـوة، وتحققـت بالتـالي النـواتج التعليميـة الـتي تسـتهدفها برامـج التعليـم، وسـتؤازر عمليـاتُ القيـاس والتقييـم والتقويـم عمليَّـتَيْ التعلـم والتعليـم.

هذه الطبعة إهـداء من المجمـع
ولا يُسمح بنشرها ورقياً أو تداولها تجارياً

أهمية الدراسة

من الممكن أن:

أ. تسـهمَ الدراسـة في تقديـم تغذيـة راجعـة للمعنيـين بِتبـنّي الاختبـارات الموضوعيـة في برامـج تعليـم العربيـة للناطقـين بلغـات أخـرى.

ب. تسـهمَ في إثـراء الأعمـال والدراسـات الـتي تعُـنى باختبـارات اللغـة في ميـدان علـم اللغـة التطبيقـي وتعليـم العربيـة للناطقـين بلغـات أخـرى عمومًـا، والاختبـارات الموضوعيـة خصوصًـا.

ج. تفيدَ كلًّا من: مديري ومخططي برامج تعليم العربية لأهلها ولغير أهلها

د. تفيـدَ الذيـن لـم ينالـوا حظًّـا مـن التدريـب في كيفيـة تصميـم الاختبـارات وإعدادهـا وبنائهـا وتقنينهـا، ومسـاعدتهم في تحديـد النـواتج التعليميـة والمسـتويات المعرفيـة.

هـ. تفيـدَ المؤسسـات التعليميـة والبرامـج الـتي تُعـنى بتعليـم العربيـة للناطقـين بلغـات أخـرى داخـل الوطـن العـربي وخارجـه.

و. تفيـدَ معلـمي اللغـات الثانيـة والأجنبيـة الذيـن يعـدون اختباراتهـم الصفيـة وغـير الصفيـة بأسـلوب يفتقـر إلى المعايـير العلميـة.

أهداف الدراسة

تهدف الدراسة من خلال تناول هذه الاختبارات إلى:

أ. إلقـاء الضـوء علـى أكـثر الاختبـارات الموضوعيـة اسـتخدامًا في السلاسـل الـتي تُعـنى بتعليـم العربيـة للناطقـين بلغـات أخـرى.

ب. بيان المستويات المعرفية التي تقيسها الاختبارات وتلك التي لا تقيسها.

ج. التعرف على النواتج التعليمية للبنود المستخدمة، والوقوف على أوجه القصور فيها.

د. إثراء الدراسات اللغوية في مجال تعليم العربية للناطقين بلغات أخرى، وفي مجال القياس والتقييم اللغوي.

هـ. فتح المجال أمام دراسات أخرى تعضد ما توصلت إليه الدراسة الحالية أو ترفضه.

أسئلة الدراسة

تحاول هذه الدراسة، في ضوء الأهداف المتوخاة، أن تجيب عما يلي:

أولاً: ما أكثر بنود الاختبارات الموضوعية استخدامًا في سلاسل تعليم العربية للناطقين بلغات أخرى؟

ثانيًا: ما الأهداف المعرفية التي تشيع في الاختبارات الموضوعية وتلك التي لا تقيسها؟

ثالثاً: ما النواتج التعليمية التي ينبغي أن تحققها الاختبارات الموضوعية والعيوب الكامنة فيها؟

حدود الدراسة

أولًا: اقتصرت الدراسة في جمع مادتها من اختبارات محددة، طبقت ولا يزال يطبق بعضها في سلاسل تبنتها مؤسسات عريقة تُعنى بتعليم العربية للناطقين بلغات أخرى حول العالم.

ثانيًا: اقتصرت الدراسة على الاختبارات الموضوعية ليس إلا، ولم تتناول أيًّا من الاختبارات (الذاتية) ولا تلك القائمة على تنفيذ المهام.

ثالثًا: اقتصر جمع المادة على الاختبارات الموضوعية التي طبقت في الفترة من ٢٠١٦- إلى قبيل الربع الأخير من عام ٢٠٢٢.

رابعًا: اقتصرت الدراسة على المجال المعرفي، بمستوياته الستة وفقًا لتصنيف بلوم، ولم تتجاوزه إلى المجالين الوجداني والنفس -حركي. ولم تعتمد تصنيف بلوم المعدل.

الإطار النظري والدراسات السابقة

أولًا: الإطار النظري:

المطلع على الأعمال التي تتناول الاختبارات الموضوعية، يجد أن دربة بعضٍ ممَّن يتصدون لها، يعتريها قصور وقلة خبرةٍ، إلا أن استخدامها في اضطراد مستمر. كما يلاحظ أيضًا أن ما تتضمنه الأدبيات المختلفة حاليًا، عن كيفية كتابة بنود الاختبار الفعالة لا يعدو أن يكون: افعل ولا تفعل، وقل ولا تقل. لذا فإن كثيرًا من العمل يكون خبط عشواء، وبالتالي قلما تعكس نواتج هذه الاختبارات الأهداف التعليمية للبرامج أو قلما ترتبط بالعملية التعليمية.

صنف بلوم الأهداف التربوية في المجال المعرفي إلى ستة أقسام رئيسة هي: التذكر، والفهم، والتطبيق، والتحليل، والتركيب، والتقويم. وقسم كلًّا منها إلى أقسام أخرى فرعية. وقد بينت الدراسات التربوية التحليلية التي بحثت في المستويات الهرمية لهذا التصنيف، أنه يركز على الأهداف التي تنتمي إلى القسم الأدنى المتعلق بالمعرفة، على الرغم من أن الأهداف التي تليه تُعدُّ أكثر ضرورة إذ هي أهداف عليا ينبغي تضمينها في تصميم المناهج وإعداد الاختبارات وجداول المواصفات. وممن أشار إلى ذلك علام (٢٠٠٤) الذي يقول بِـ« أن التوجه الجديد الآن هو أن التطبيق والفهم مترادفان» (ص.٨٩). وكان ستيقينز Stiggins (٢٠٠١، كما ورد في علام،٢٠٠٤) قد قدم أدلة تجريبية تدحض فرضية البنية الهرمية

للمستويات. وفي العام نفسه، قام كراثول Krathwohl (٢٠٠١، كما ورد في علام أيضًا، ٢٠٠٤) «بإعادة النظر في تصنيف بلوم ومراجعته استنادًا إلى النتائج المتجمعة للبحوث والدراسات السابقة» (ص. ٨٩).

ينبغي أن تعكس نواتج التعلم التي يقيسها أيُّ اختبار-بأمانة- الأهداف التعليمية المنصوص عليها في البرنامج. وهكذا فإن أول عمل ينبغي الاهتمام به هو تحديد هذه الأهداف، ثم التحقق بأنها قد صيغت في عبارات يسهل فهمها لتُعَدَّ البنود في ضوئها. إلا أن هذا الكلام سهل القول، صعب التنفيذ. ويكون الأمر أكثر صعوبة إذا لم تكن هناك أهداف تعلمية وتعليمية واضحة محددة المعالم منصوص عليها في البرامج التعليمية.

يشير فضل (٢٠٠٧) إلى أنه بعد أن تحدد نواتج التعلم «وتوصف محتويات المادة التي ستدرس في أيٍّ من برامج اللغة، ينبغي إعداد قوائم أو جداول بالمواصفات، وهذه القوائم من أوجب مهامها أن تربط النواتج بالمحتوى، وأن تشير إلى الوزن أو الثقل النسبي الذي يعطى لمختلف الجوانب في المادة» (ص. ١٢٠). وتأتي أهمية قوائم أو جداول المواصفات، كما يذكر كل من الدامغ وفضل (٢٠١٢) من أنها «تساعد القائمين على أمر الاختبار في:

١. الحكم الدقيق- نسبيًّا- على مدى صلاحية الاختبار.

٢. الحكم الدقيق- نسبيًّا- على مدى تحصيل الطلاب.

٣. التزويد بدليل أو مرشد واضح لقياس الأهداف المنشودة.

يكون هذا الجدول ثنائي الأبعاد؛ حيث يمثل البعد الأفقي منه موضوعات المحتوى التي يتناولها الاختبار، ويمثل البعد الرأسي الأهداف. ويقوم من يعد الاختبار، بتقدير النسبة المئوية (الثِّقل) لكل موضوع في خانة المحتويات؛ وفق الثقل الذي يمثله في الجزء المقرر من المنهج. ومن الممكن الاستفادة من هذه النسب والثِّقل «من وجهين:

(أ) تحديد عدد بنود المحتوى المرتبطة بالصيغ السلوكية المخصصة لكل مستوى من مستويات الأهداف.

(ب) تحديد عدد البنود التي يجب أن تخصص لكل موضوع من موضوعات المحتوى في كل مستوى من مستويات الأهداف من بنود الاختبار ككل» (ص.٢٦).

فعلى سبيل المثال، إذا ما أراد معلم في برنامج اللغة أن يضع جدولًا للمواصفات ليساعده في اختبار للقواعد، ورأى أن يقيس بعض الأهداف الموضحة من منظومة الأهداف: الفهم-التطبيق -التحليل، فقام بتوزيع الموضوعات في ست مجموعات في مجالي الأسماء والأفعال كالتالي:

أ. مجموعة الأسماء المبنية. ب. مجموعة الأسماء المعربة. ج. مجموعة الأسماء الممنوعة من الصرف.

د. مجموعة الأفعال المبنية. هـ. مجموعة الأفعال المعربة. و. مجموعة الأفعال المعتلة.

جدول رقم (١) نموذج لقائمة بمواصفات اختبار في قواعد اللغة العربية

المحتوى الهدف	الاسم المبني	الاسم المعرب	الممنوع من الصرف	الفعل المبني	الفعل المعرب	الفعل المعتل	المجموع
الفهم	٦	٤	٤	٢	٢	٤	٢٢
التطبيق	٦	٨	٦	٤	٨	٦	٣٨
التحليل	١٠	١٢	٦	٤	٤	٤	٤٠
المجموع	٢٢	٢٤	١٦	١٠	١٤	١٤	١٠٠

ثانيًا: الدراسات السابقة

دراسات تناولت الشيوع

توجد دراسات تتَبَّعت أوجه القصور من حيث: شيوع أنواع بعينها من البنود، ودراسات أخرى تقصَّت المستويات المعرفية التي تشيع في هذه الاختبارات، وهناك من الدراسات والأدبيات ما بحث في النواتج التعليمية للأنواع المختلفة للاختبارات الموضوعية والتي غالبًا ما تركز على مستويات: الحفظ والتذكر والاستيعاب، ولا تتعداها إلى المستويات العليا كالتحليل، والتقويم.

تتفاوت الدراسات السابقة في مدى صلتها بموضوع الدراسة؛ فبعضها علاقته عامة، وبعضها الآخر علاقته جزئية؛ بمعنى أن جزءًا أو مكونًا منها ذو علاقة بمتغيراتها، إلا أن الاستفادة من كلا النوعيْن كانت واضحة؛ وأُثبِت كل ذلك في مكانه في ثنايا العمل. وستحاول الدراسة عرض جوانب من تلك الدراسات: الموافق أو المناقض لها.

ففي دراسة أعدها عثمان (٢٠١٥) أوضحت النتائج «اعتماد أسئلة الاختيار من متعدد بنسبة ١٠٠٪، وقد كانت نسبة التذكر منها ٣٠٪، والفهم ٤٠٪، والتطبيق ٣٠٪، ولم تُستخدمْ فيها مستويات: التحليل والتقويم والإبداع». (ص.١٧). وأظهرت نتائج دراسة الزبون (٢٠١٣) أن هناك فروقًا «ذات دلالة إحصائية في كفاياتهم [المعلمين] المتعلقة ببناء الاختبارات التحصيلية تبعا لمتغير المؤهل العلمي، والدورات التأهيلية، ودراسة مساق الاختبارات والمقاييس، وأن أكثر أنواع أدوات التقييم استخدامًا كان اختبارات الاختيار من متعدد، يليها اختبارات الصواب والخطأ، ثم اختبارات المزاوجة». كما توصلت إلى أن هناك معيقات تحول دون بناء الاختبارات التحصيلية وفق معايير الاختبار الجيد. «تركزت على عدم المعرفة بطرق تفسير النتائج... وعدم المعرفة الكافية بكيفية بناء جدول المواصفات، وصعوبة تحليل محتوى المادة» (ص. ٩٣-٩٤).

في دراسة دروزة (٢٠١١) عن مدى مراعاة مستويات بلوم للأهداف المعرفية، والمستويات تستخدم أكثر من غيرها، كان «ما توصلت إليه من نتائج ينسجم مع الترتيب الهرمي الذي وضعه بلوم؛ حيث كانت أعلى نسبة لمستوى التذكر يليها الفهم، فالتحليل، فالتطبيق، فالتركيب...، فالتقويم» (ص.٢٥٧٧).

وفي الدراسة التي أعدتها الصرايرة (٢٠١١) جاءت الأسئلة في مستوى التذكر بالمرتبة الأولى بنسبة ٦٦,٥٥٪ في حين جاءت الأسئلة التي تستدعي مستوى الفهم والاستيعاب بالمرتبة الثانية بنسبة ٢٢,٣٣٪، والأسئلة في بقية المستويات بنسبة ٨,٢٢٪؛ أي بانخفاض بنسبة الأسئلة المتعلقة بالعمليات العقلية العليا، وقد عزت السبب إلى المعلمين لأن «برامج إعدادهم قبل الخدمة قد لا تكون بالمستوى المعرفي الجيد لإعداد مثل هذه الاختبارات حسب المستويات العليا؛ [لذا] توصي بعقد دورات متخصصة سواء قبل أو بعد [هكذا] الخدمة في إعداد الاختبارات الكتابية التحصيلية وفق معايير محددة» (ص٥٢).

توصلت دراسة الجواد (٢٠١١) إلى أن أسئلة الفهم قد احتلت المرتبة الأولى، يليها على الترتيب أسئلة التذكر، وأسئلة التركيب، وأسئلة التحليل، فأسئلة التطبيق، ثم أسئلة التقويم.

وأظهرت نتائج دراسة الجعافرة (٢٠٠٩) أن غالبية الأسئلة ركزت على التذكر والفهم، «في حين كانت الأسئلة التي تقيس أهداف التطبيق متوسطة، في حين كانت الأسئلة في المستويات العليا قليلة. كما أظهرت النتائج بأن الأسئلة الموضوعية كانت في معظمها من نوع أسئلة التكميل بنسبة بلغت ٧٠,٣٪، وسبب ذلك سهولة وسرعة إعدادها» (ص.٧٧).

وفي الدراسة التي قام بها فضل (٢٠٠٧) يذكر «أن أكثر أنواع بنود الاختبارات الموضوعية شيوعًا هي بنود الاختيار من متعدد؛ فقد احتفظت بالمرتبة الأولى من حيث كثرة الاستخدام» (ص.١٩٧). و«بأن نسبة تبلغ نحو ٩٥٪ من

المستطلَعين أقرت بوجود القصور في الاختبارات الموضوعية التي تقيس عناصر اللغة ومهاراتها» (١٩٨). وفي الدراسة نفسها استنتج أن قياس مستوى الفهم قد احتفظ بالرتبة الأولى سواءً في اختبارات البرامج أو السلاسل أو متوسط الاستخدام بنسب تجاوزت النصف بقليل.

تشير دراسة هاشم (د.ت) أن أسئلة الاختيار من متعدد كانت أكثر الأسئلة الموضوعية شيوعًا، وبأنها «تقيس نواتج التعلم بكفاءة... ويمكن استخدامها في قياس مهارات التفكير العليا على نحو جيد». (ص. ١٢).

ومن تحليل البيانات في دراسة الدامغ وسكُلْفيلد (Addamigh & Scholfield 2005) تبيَّن أنه كان «على رأس قائمة الاستخدام بنود الاختبارات الموضوعية وفي مقدمتها الاختيار من متعدد بنسبة ٣٠,٧٪. أما بقية الأنواع وعددها أربعة وثلاثون نوعًا، فقد تشاركت في نسبة ٦٩,٣٪ المتبقية. وجاء نوع تكملة الفراغ في المرتبة الثانية، وبنود المزاوجة في الدرجة الثالثة» (ص.١١).

استنتجت دراسة السليطي (٢٠٠٤) أن ثمة عدم توازن في توزيع الأسئلة على المستويات المعرفية المختلفة، «حيث تركز بنسبة كبيرة على مستوى التذكر أكثر من باقي المستويات. وتجب مراعاة أن تكون أسئلة الاختبار شاملة لجميع مستويات الفهم في التفكير، وألا تقتصر على الفهم المباشر أو الفهم الاستنتاجي» (ص. ٩٧).

توصل عبد الرازق (٢٠٠٢) إلى أن تدريب عينة الدراسة – من خلال الورش التعليمية – على إنتاج أسئلة تحريرية تقيس مستويات معرفية إدراكية مختلفة، وتتفق وقواعد صياغة السؤال الجيد، أدى إلى تحسن مهاراتهم في صياغة أسئلة تحريرية تقيس مستويات معرفية إدراكية مختلفة.

استنتجت الدراسة التي قام بها حطاب (٢٠٠٥) إلى أن الاختبارات الموضوعية «ركزت على المستويات المعرفية الدنيا من تصنيف بلوم للمجال المعرفي وبخاصة

مستوى التذكر، وكان الإهمال واضحًا للمستويات العليا، وبخاصة مستويا التركيب والتقويم» (ص.٥١).

دراسات تناولت النواتج:

تفرق الدراسات والأدبيات التي تناولت النواتج التعليمية، بين ما يجب أن يعرفه الطالب من قبيل: المبادئ والقواعد والأسس والنظريات والمفاهيم... إلخ، وما يجب أن يقوم بأدائه من مهارات عقلية وعملية عامة. ويتمثل أهم ناتج من نواتج التعلم، فيما يخص تعلم اللغات، في زيادة معدل الأداء اللغوي في المستويات العليا للتفكير. وتُعنى الدراسة الحالية بالمجال المعرفي الذي يتضمن الأنشطة الفكرية والمعرفية، علاوة على القدرات العقلية لدى المتعلم، ويتلخص هذا المجال في مدى مقدرة الفرد على حفظ المعلومات واستدعائها عند الحاجة، بالإضافة إلى مدى قابلية الفرد على: الفهم والتحليل والتطبيق بشكل عام.

تساعد الصياغة الدقيقة لنواتج التعلم على مساعدة المعلم والمتعلم على اكتساب نواتج التعلم المقصودة، ثم اختيار أساليب التقويم الموضوعية المناسبة. ويلخص عبده (٢٠٢١) أهمية نواتج التعلم للطالب في نقاط أهمها: (أ) معرفة ما الذي سيكتسبه ويوجه نشاطه ودافعيته على تحقيق تلك النواتج. (ب) توجيهه للتعلم في ضوء أهداف محددة وواضحة (جـ) توفير فرصة له لتحمل مسؤولية أكبر من خلال التعلم الذاتي. (د) المساعدة على التقويم الذاتي وتطوير الأداء أولًا بأول في ضوء قواعد واضحة ومحددة.

من الأدبيات التي بحثت في نواتج التعلم، دراسة أبو عيش (٢٠١٦) التي وجدت أن غالبية مؤشرات أساليب تقويم نواتج التعلم تقع بين المتوسط والمرتفع. وأوصت «بعدم الاقتصار على أساليب التقويم التقليدية، وإعداد دورات تنمية مهنية لأعضاء هيئة التدريس في مجال القياس والتقويم بحيث تتناول أساليب التقويم الحديثة، وممارسات تقويمية يمكن استخدامها في تقويم نواتج التعلم،

وخاصة تلك التي تدعم تنمية التفكير الإبداعي، والتفكير الناقد» (ص. ٦٢).

جاء في دراسة هاشم وآخرين (د.ت) أنه كان من أهم نقاط الضعف الخاصة بطرائق التقويم «أن الامتحانات لم تكن تتصف بالموضوعية، ولم تغط محتويات المقرر... وغلب عليها استدعاء المعرفة» (ص.٤). وقد أوصت تلك الدراسة، بأحد عشر عنصرًا تنبغي مراعاتها في عمليات التقويم، كان آخرها ذكرًا، أنه ينبغي «اتساق أسلوب التقويم مع نواتج التعلم... حيث تختلف أساليب التقويم الخاصة بالمهارات عن تلك الخاصة بتحصيل المعارف» (ص.٢٤).

وفي دراسة أعدها آل داود وآخرون (٢٠١٦) أشارت إلى أن تحديد نواتج التعلم، من شأنه أن يؤدي إلى معرفة طرائق التدريس المناسبة، واختيار أساليب التقويم الحديثة للتحقق من اكتساب المتعلمين لنواتج التعلم المقصودة. وتذكر الدراسة نفسها كما ورد في الشدوح (٢٠٠٩) بأن إجراءات تقويم نواتج التعلم التي بحثوا فيها «لا تقوم على معايير علمية واضحة، وإنما تستند على خيارات ورؤى الأساتذة، دون العمل وفق معايير وأطر موحدة، كما أن التقويم لا يقيس ما ينبغي تحصيله من مهارات ومعارف بحسب أدلة الجامعات والكليات التي اطلعوا عليها» (ص.١٧١).

وتوجز الدراسة- فيما يلي- النواتج التي يُفترض أن تقيسها الاختبارات الموضوعية كما وردت في بعض الأدبيات والدراسات، مع ذكر عيوب كل منها.

أولًا: الاختيار من متعدد: يقول قرونلند Gronlund(1990) إن هذه البنود تستخدم لِقياسِ:

١. «نواتج التذكر مثل: معرفة المصطلحات، والعادات، والاتجاهات، والتصنيفات، والمبادئ والتعميمات».

٢. النواتج عند مستويَيْ الفهم والتطبيق مثل: القدرة على تحديد الحقائق والأساسيات وتطبيقها، وتفسير العلاقة بين الأسباب والنتائج.

٣. النواتج عند مستويات التحليل والتركيب والتقويم مثل: القدرة على تحليل عناصر المواقف الجديدة إلى عناصرها الجزئية من خلال استخدام معايير معينة لتحليل هذه المواقف» (ص.٤٥).

العيوب: من العيوب التي أخذت على بنود الاختيار من متعدد أنها:

١. إن لم تعد إعدادًا متقنًا، فإنها سوف تقتصر على قياس التذكر والفهم.

٢. تفسح المجال للغش والتخمين.

٣. لا تقيس نواتج التعلم التي تتعلق بالتأليف والتنظيم والابتكار.

٤. قد تعطي انطباعًا للطالب بأن كل المشكلات، ليس لديها إلا حل واحد ووحيد.

٥. قد لا يختار الطلاب النابهون الجواب الصحيح؛ لغموض في الصياغة، أو لوجهات النظر المختلفة، أو أن لديهم مزيدًا من المعارف عن الموضوع لا تكون عند بقية الطلاب.

٦. يمكن أن تكون مصوغة صياغة ركيكة جيدة ظاهريًا، إلا أنها تقيس أشياء هامشية تقتصر على معرفة الحقائق.

٧. باستعمال بنود الاختيار من متعدد، لا يعطى الطلاب الفرصة للكتابة عن الموضوع الذي يتعلمونه؛ لأن

٨. عليهم الاختيار من قائمة محددة من الخيارات؛ مما يحول بينهم وبين التعبير عن إبداعهم.

ثانيًا: الصواب والخطأ: يذكر عبد الرازق (٢٠٠٢) فيما يخص النواتج التعليمية أن هذه البنود لا تصلح في معظم الأحيان إلا لقياس التذكر، ويُستثنى من ذلك

«قياس معرفة الحقيقة من الرأي، أو قياس علاقات السبب والنتيجة، ولذلك فإن كثيرًا من المخرجات التي يمكن قياسها بوساطة هذا النوع من الأسئلة، يمكن قياسها بكفاءة أكبر باستخدام الأنواع الأخرى من أسئلة التعرف، مثل الاختيار من متعدد».

العيوب: يوجه لهذا النوع كثير من النقد. ويدور معظمه، كما يشير الدامغ وفضل (٢٠١٢) حول أمرين، فإنها:

١. تعتمد -إلى حد كبير- على استظهار حقائق منفصلة غير مترابطة؛ وهذا من شأنه أن يُتَفّهَ من أهمية فهم هذه الحقائق.

٢. تفتقر للثبات؛ لأن فرصة الوصول للجواب الصحيح بوساطة التخمين تصل إلى نسبة ٥٠٪ (ص.٧٧).

ثالثًا: المزاوجة: يشير قرونلند (Gronlund 1990) بأن بند المزاوجة ما هو إلا صورة معدلة من صور الاختيار من متعدد؛ فبدلًا من أن توضع الاستجابات المحتملة في قائمة تحت كل ساق، فإن سلسلة من السوق (المقدمات المنطقية) توضع في قائمة واحدة، وتوضع الاستجابات في قائمة أخرى (ص.٤٥).

وفي ضوء آراء قرونلند أعلاه، فإن النواتج التعليمية التي تقيسها بنود المزاوجة تكون شبيهة – لحد ما – بالنواتج التي تقيسها بنود الاختيار من متعدد.

العيوب: يمكن أن توجز العيوب لهذا النوع في:

١. أنها تقيس المعلومات عن الحقائق التي يستظهرها الطلاب في أدنى سلم المستويات. إلا أن أوسترلند (Osterlind 1992) رد على هذا النقد بأن «هذا النوع من البنود يمكن أن يجبر المفحوص على استخدام مهارات تفكير معقدة في بنود معينة» (ص.٢٤). ويبرهن على كلامه ببعض الأمثلة التي استدل بها.

٢. يحتاج هذا النوع إلى وجود عدد كبير من العلاقات المتناظرة أو المترابطة في المادة المفحوصة؛ وقد لا يتوافر ذلك دائمًا.

٣. لا تقيس الفهم والقدرة على التمييز بنجاح.

رابعًا: إعادة الترتيب: ترد هذه البنود على هيئة حروف متناثرة لتصير كلمة، أو كلمات لإعادة ترتيبها لتصير جملةً، أو جمل لتكوِّن فقرة، أو عدة فقرات لتكوِّن موضوعًا أو قصة. ويصلح هذا النوع من البنود في اختبارات عناصر اللغة؛ مثل اختبارات المفردات أو اختبارات القواعد والتراكيب، كما تصلح في اختبارات التعبير الموجه للصفوف المبتدئة.

العيوب: على الرغم من أن هذا النوع من البنود يكون عامل التخمين فيه قليلًا جدًّا، إلا أنه يستغرق وقتًا طويلًا في إعداده، كما أن خطأ الطالب في جزئية منه يترتب عليه خطأ في جزئية أخرى تليه، أو كلها أحيانًا.

خامساً: التكميل: يعد هذا النوع وسطًا بين الاختبارات الموضوعية والاختبارات الذاتية؛ فإذا أحسن إعداده بحيث يكون المطلوب كلمة واحدة أو عبارة واحدة ووحيدة، فإنها تكون بذلك موضوعية مئة في المئة. أما إذا كانت التكملة بأي كلمة صحيحة غير محددة، فإن التكميل في - هذه الحالة - ينتمي إلى الاختبارات ذاتية التصحيح.

يلخص سمارة (٢٠٠٢) نواتج التعلم التي تقيسها بنود التكملة فيما يلي:

١. قياس النواتج المتعلقة بالحقائق البسيطة كالأسماء، والتواريخ، والأماكن، والأحداث.

٢. قياس النواتج المتعلقة بمعرفة المبادئ.

٣. اختبار معاني المفردات والمصطلحات.

٤. قياس النواتج المتعلقة بمعرفة الطريقة أو الأسلوب (ص.٨٨).

سادسًا: الأسئلة ذات الإجابات القصيرة: يشابه هذا النوع -إلى حد كبير- بنود التكميل في بعض الأوجه، كما أنه يعد أيضًا حلقة وصل بين الاختبارات الموضوعية والاختبارات الذاتية. وكلاهما ينتمي إلى ما كان يطلق عليه - حتى وقت قريب - بنود ملء الفراغ، ولكن هذه التسمية يتحاشاها خبراء القياس والتقويم كما يقول أوسترلند Osterlind (١٩٩٢).

تقيس عادة، وبصورة رئيسة، النواتج المتعلقة باستدعاء الحقائق؛ حيث إن تصميمها وشكلها لا يسمح سوى بتذكر المعلومات أو البيانات التي يمكن كتابتها في كلمة واحدة، أو كلمتين، أو في عبارة قصيرة واحدة أو كتابة عدد، أو رمز، أو معاني المفردات. وفي هذا الصدد تعد أكثر تقييدًا عند تطبيقها في اختبارات التحصيل، والاستعداد أو القدرة من بنود الاختيار من متعدد، والصواب والخطأ أو حتى المزاوجة.

العيوب: من العيوب التي توجه إلى هذه الأسئلة:

١. أنها لا تخلو من صفة الذاتية في التصحيح؛ فلا يستطيع المعلم توقع كل الاستجابات التي يدلي بها الطلاب.

٢. إن تقييد الطالب بالإجابة بكلمات قليلة، لن يقيس سوى تذكّر حقائق محددة وأسماء، وأماكنَ وأحداث، على حساب سلوكيات أكثر تعقيدًا.

٣. قد تكون هناك أخطاء ليست في الحقائق، وإنما أخطاء في التهجئة الإملائية أو النحو تعقد قضية التصحيح.

٤. إذا حدثت الأخطاء أعلاه، فإنها تبطئ من عملية التصحيح.

عزت بعض الدراسات، مثل: (الزيون ٢٠١٣، والصراصرة٢٠١١، والجعافرة ٢٠٠٩، وفضل ٢٠٠٧) كثيرًا من أسباب القصور في عدم تحقيق النواتج التعليمية للاختبارات الموضوعية إلى كل أو بعضٍ مما يلي:

١. قلة أو عدم خبرة المعلمين الذين يستخدمون هذا النوع من الاختبارات.

٢. عدم تلقي المعلمين دورات وتدريبات كافية على تصميمها وبنائها وإعدادها.

٣. عدم القدرة على تحديد الأهداف والنواتج التعلمية والتعليمية.

٤. عدم المعرفة الكافية بكيفية بناء جداول المواصفات.

٥. صعوبة بناء هذه الاختبارات، ولأنها تستغرق وقتًا طويلًا في التصميم والبناء. وصعوبة التقيد بشروطها وضوابطها.

٦. عدم التقيد بالمعايير العلمية عند إعداد هذه الاختبارات.

٧. عدمِ تضمين البرامج اختبارات موضوعية نموذجية يستأنس بها معدو الاختبارات.

منهج الدراسة

اعتمدت الدراسة منهجًا وصفيًّا تحليليًّا؛ معتمدة على الأسلوب الكمي والأسلوب الكيفي في تحليل المادة التي تجمعت لديها. وللوصول إلى ذلك نظرت في كل ما توافر لها من: أدبيات، ودراسات، وأعمال تناولت الاختبارات الموضوعية في اللغتين العربية والإنجليزية، إضافة إلى ما لدى الدارس من خبرات متواضعة في المجال.

بعد عرض أداة الدراسة والتحليل على فريق من ذوي الخبرة، بغرض التثبت من مدى صلاحيتها، حللت ما توافر لديها من بيانات. وباتباع هذا النهج، تمكنت الدراسة من تحليل تلك الاختبارات وتقييمها لعرض ما يعتريها من قصور؛ ثم توصلت إلى نتائج واقتراحات قد يكون من شأنها تحسين هذه الاختبارات وتطويرها لتقيس المستويات والنواتج المستهدفة.

مجتمع الدراسة والعينة

تألف مجتمع الدراسة من عدد كبير من اختبارات جمعت من سلاسل في تعليم اللغة العربية للناطقين بلغات أخرى، تستخدم في جامعات ومؤسسات رائدة حول العالم، أعدها مؤلفون من ذوي الخبرة في إعداد المادة اللغوية، إلا أن من بينهم من ليست لديه خبرة أو دربة في مجال اختبارات اللغة. أما عينة الدراسة، فهي نفسها مجتمع الدراسة المتمثل في هذه الاختبارات التي تراوحت بين: تحصيلية، وتحديد مستوى، وتشخيصية.

بلغ عدد البنود والأسئلة التي حللت ٢,٩٣٣ بندًا. تقصت الدراسة فيها البنود الموضوعية في اختبارات العناصر والمهارات التالية: (أ) المفردات (الألفاظ) و(ب) التراكيب والقواعد (النحو والصرف) و(جـ) مهارة الاستماع (فهم المسموع) و (د) مهارة القراءة (فهم المقروء) و (هـ) الكتابة (للمبتدئين). وهذه السلاسل هي:

- العربية للعالم، معهد اللغويات العربية بجامعة الملك سعود، سلسلة من منشورات دار نشر جامعة الملك سعود.
- عربية القرآن، سلسلة من منشورات العربية بلا حدود.
- تعلم العربية- للمرحلة الثانوية، سلسلة من منشورات مؤسسة مناهج العالمية.
- العربية بين يديك، الإصدار الثاني، سلسلة من منشورات شركة العربية للجميع.
- كنوز، سلسلة من منشورات معهد عربي للغة العربية.

أشير لهذه السلاسل في متن العمل؛ لأغراض هذه الدراسة وبهدف التيسير، وَفْقَ التسلسل التالي:

١. العربية للعالم.

٢. عربية القرآن.

٣. تعلم العربية.

٤. العربية بين يديك.

٥. كنوز.

أداة الدراسة وطريقة التحليل

(أ) أداة الدراسة

آثرت الدراسة استخدام أداة تساعد على جمع البيانات المستهدفة بطريقة دقيقة وسريعة، فاختارت لذلك بطاقات التحليل. وتألف النوع الأول لهذه البطاقات من ست بطاقات، صُممت لتجميع أنواع البنود والأسئلة ثم تحليلها للوقوف على أكثرها استخدامًا في السلاسل قيد الدراسة، وجاءت كالتالي: بطاقات تحليل بنود...

١. الاختيار من متعدد.

٢. التكملة.

٣. المزاوجة.

٤. الإجابات القصيرة.

٥. الصواب والخطأ.

٦. إعادة الترتيب.

أما النوع الثاني من بطاقات التحليل، فهو الخاص بمدى شيوع المستويات المعرفية، فقد احتوت على مستوى الهدف المعرفي الذي يقيسه كل بند من

البنود الموضوعية في ضوء المستويات المعرفية الستة في تصنيف بلوم وهي: التذكر-الفهم-التطبيق-التحليل-التركيب-التقويم. ورُتِّبت وَفْقَ شيوع هذه المستويات في السلاسل قيد الدراسة. وكان الاعتماد عند تصميم هذه البطاقات على: مصادر المعرفة، والدراسات السابقة، ودراسات تناولت تحليل الكتب والاختبارات وتقويمها، بالإضافة إلى استشارة ذوي الخبرة وآراء المحكمين.

صدق الأداة وثباتها

- الصدق: للتحقق من صدق أداة الدراسة، عرضت على عدد من المتخصصين في القياس والتقويم، بالإضافة إلى متخصصين في المناهج وطرائق التدريس؛ للتأكد من الصدق الظاهري للأداة، وللتثبت من مدى صدق محاور الأنواع والمستويات التي تتناولها، وصلاحيتها لجمع البيانات.
- الثبات: أعيد تحليل المادة قيد الدراسة، بصورة مستقلة عن الباحث؛ للتأكد من ثبات التحليل. وبمقارنة نتائج تحليل المضمون، اتضح أن هناك اتفاقًا كبيرًا بين نتائج التحليلين؛ وبذا أمكن الاطمئنان إلى أن عمليات التحليل تفي بأغراض هذه الدراسة، وأنها يمكن الوثوق فيها إلى درجة مقبولة.

(ب) طريقة التحليل

اختيرت طريقة تحليل المحتوى؛ لأغراض التصنيف والتلخيص والمعالجة الإحصائية؛ لأنها تجمع بين ميزات طرائق الدراسات الأخرى من حيث: الموضوعية، والاتساق، والتعميم، والتكميم. وكما يشير بركات (١٩٩٤)

فمن حيث «الموضوعية؛ فالتحليل يستند إلى أسس وقواعد ثابتة، تمكّن الباحثين المختلفين من الوصول إلى النتائج نفسها من الوثائق الموحدة. ومن حيث الاتساق؛ لأن التحليل يستند إلى تنظيم ثابت ومعيار محدد يمكن أن تبوب على أساسه كل وحدات التحليل. وأما من حيث العمومية؛ فإن هذه الطريقة ذات

هذه الطبعة إهـداء من المجمع
ولا يُسمح بنشرها ورقياً أو تداولها تجارياً

قابلية للتعميم، بأن يكون هناك مجال للإفادة من نتائج التحليل بصورة عامة. وأما التكميم، فيعد ميزة خاصة في طريقة تحليل المحتوى، لا تشترك معها فيها بقية طرائق الدراسات البحثية، وهذا يعني أنه لا معنى لطريقة تحليل المحتوى إذا لم تؤدِّ إلى نتائج كمية يمكن معالجتها إحصائيًا» (ص.٢٠٧).

عرض النتائج ومناقشتها

ستحاول الدراسة في هذا القسم الإجابة عن أسئلة الدراسة، من خلال عرض النتائج التي استقتها من بطاقات التحليل؛ الخاصة بأنواع البنود ومستويات المجال المعرفي.

فللإجابة عن السؤال الأول، ونصه: ما أكثر بنود الاختبارات الموضوعية استخدامًا في سلاسل تعليم العربية للناطقين بلغات أخرى؟

استقصت الدارسة الاختبارات الموضوعية المستخدمة في العينة للوقوف على مدى شيوع الاستخدام، وجاءت البيانات كما يوضحها الجدول رقم (١).

الجدول رقم (١) عدد البنود المستخدمة في السلاسل ونسبها المئوية.

نوع البند / السلسلة	الاختيار من متعدد	الصواب والخطأ	المزاوجة	إعادة الترتيب	التكملة	الإجابات القصيرة	المجموع
١. **العربية للعالم**							
عدد البنود	٢٢٦	٦٢	٧٧	٥٠	٥٢	١٠٨	٥٧٥
نسبة الاستخدام	٪٣٩,٣	٪١٠,٨	٪١٣,٤	٪٨,٧	٪٩	٪١٨,٨	٪١٠٠
٢. **عربية القرآن**							
عدد البنود	٣٢٣	٧٧	٢٠٨	–	٧٤	١٠٤	٧٨٦
نسبة الاستخدام	٪٤٧	٪١١	٪٣٠		٪١٠	٪١٥	٪١٠٠
٣. **تعلم العربية**							
عدد البنود	٢٠١	٢٥	١٢	–	١٠	–	٢٤٨
نسبة الاستخدام	٪٨١	٪١٠,٢	٪٤,٨٥		٪٤		٪١٠٠
٤. **العربية بين يديك**							
عدد البنود	٢٠٧	١٠٨	١٠٥	٣٧	١٩٠	١٨	٦٦٥
نسبة الاستخدام	٪٣١,١	٪١٦,٢٤	٪١٥,٨	٪٥,٥٦	٪٢٨,٦	٪٢,٧	٪١٠٠
٥. **كنوز**							
عدد البنود	١٣٤	٢٨	٧٦	١٩	١٧٣	٢٢٩	٦٥٩
نسبة الاستخدام	٪٢٠,٣	٪٤,٣	٪١١,٥	٪٢,٩	٪٢٦,٣	٪٣٤,٧	٪١٠٠
المجموع الكلي للبنود =	١,٠٩١	٣٠٠	٤٧٨	١٠٦	٤٩٩	٤٥٩	٢,٩٣٣

يتضح من الجدول، أن البنود التي يشيع استخدامها في هذه السلاسل كالتالي:

١. بنود الاختيار من متعدد: وحظيت بالعدد الأكثر؛ إذ استحوذت على نسبة قدرها ٣٧,٢٪، وجاء استخدامها على رأس القائمة في كل السلاسل قيد الدراسة عدا السلسلة رقم (٥) التي حلَّ فيها في المرتبة الثالثة، بعد الإجابات القصيرة والتكملة، بنسبة بلغت ٢٠,٣٪، وكانت أعلى نسبة استخدام له في السلسلة رقم (٢) بنسبة قدرها ٤١,١٪.

٢. التكملة: وحلت في المرتبة الثانية بنسبة بلغت ١٧٪. وجاء استخدامها على رأس القائمة في السلسلة رقم (٤) بنسبة تبلغ ٢٨,٦٪، وكان أقل استخدام لها في السلسلة رقم (٣) بنسبة استخدام لم تتجاوز ٤٪.

٣. المزاوجة: جاءت في المرتبة الثالثة، واستخدمت بنسبة بلغت ١٦,٣٪. وجاءت أعلى نسبة استخدام لها في السلسلة رقم (٢) بنسبة قدرها ٢٦,٥٪. أما أدنى استخدام لها، فكان في السلسلة رقم (٣) بنسبة بلغت ٤,٨٥٪.

٤. الإجابات القصيرة: جاءت في المرتبة الرابعة بنسبة قدرها ١٥,٦٪. ويتضح من الجدول أن أكثر السلاسل استخدامًا لهذا النوع، كان في السلسلة رقم (٥) بنسبة استخدام بلغت ٣٤,٧٪، أما السلسلة الثالثة، فلم تستخدم هذا النوع قط.

٥. بنود الصواب والخطأ: وحلت في المرتبة الخامسة بنسبة قدرها ١٠,٣٪. ويتضح من الجدول أن أكثر السلاسل استخدامًا لهذا النوع، كان في السلسلة رقم (٤) بنسبة استخدام بلغت ١٦,٢٤٪. أما أدنى استخدام لها، فكان في السلسلة رقم (٥) بنسبة بلغت ٤,٣٪ مما مجموعه ستمئة وخمسة وتسعين بندًا من البنود المستخدمة في السلسلة.

٦. بنود إعادة الترتيب: وحلت في ذيل القائمة بنسبة متواضعة قدرها ٣,٦٪ ليس إلا. وكانت أعلى نسبة استخدام لها في السلسلة رقم (١) بنسبة

استخدام ٨,٧٪. وسجل هذا النوع أعلى نسبة من عدم الاستخدام بين السلاسل الست؛ إذ لم يستخدم في سلسلتين منها قط.

الجدول رقم (٢) البنود الموضوعية مرتبة حسب شيوعها ونسبها.

نوع البند	المجموع العام	النسبة٪	الرتبة
الاختيار من متعدد	١,٠٩١	٣٧,٢٪	١
التكملة	٤٩٩	١٧٪	٢
المزاوجة	٤٧٨	١٦,٣٪	٣
الإجابات القصيرة	٤٥٩	١٥,٦	٤
الصواب والخطأ	٣٠٠	١٠,٣٪	٥
إعادة الترتيب	١٠٦	٣,٦٪	٦
المجموع	٢,٩٣٣	١٠٠٪	

يتضح من الجدول رقم (٢) أن بنود الاختيار من متعدد هي أكثر أنواع بنود الاختبارات الموضوعية شيوعًا؛ فقد احتفظت بالمرتبة الأولى من حيث كثرة الاستخدام. ويشير ذلك - بالتأكيد - إلى مدى أهمية هذا النوع من البنود؛ بغض النظر عن مدى تقيد معدي هذه الاختبارات بالقواعد الصحيحة لصياغتها كما سيتضح لاحقًا.

مناقشة النتائج للسؤال الأول

حظيت بنود الاختيار من متعدد بالمرتبة الأولى؛ إذ استحوذت على نسبة عامة قدرها ٣٧,٢٪، وهي نسبة قريبة من نسبة استخدامها في دراسة فضل السابقة (٢٠٠٧) الذي استنتج أن أكثر أنواع بنود الاختبارات الموضوعية شيوعًا هي بنود الاختيار من متعدد؛ فقد احتفظت بالمرتبة الأولى من حيث كثرة الاستخدام. وكذلك في دراسة أبو عيش (٢٠١٦) التي توصلت إلى أن المدرسين يكتفون بنوعين من أنواع البنود الموضوعية هما الاختيار من متعدد والصواب والخطأ بنسبة

٧٥٪ و٧٦٪ على التوالي. والاختلاف هنا أن بنود الصواب والخطأ جاءت في المرتبة الأولى، في حين أن المرتبة الأولى في الدراسة الحالية كانت من نصيب بنود الاختيار من متعدد، أما بنود الصواب والخطأ، فقد حازت على المرتبة الخامسة فيها.

وتتفق نتائج الدراسة الحالية، جزئيًا أيضًا، مع النتيجة التي توصلت إليها دراسة عثمان (٢٠١٥) التي كان من بين نتائجها أن أسئلة الاختيار من متعدد استخدمت بنسبة ١٠٠٪.

وفي نتائج دراسة الزبون (٢٠١٣) نجد أن أكثر أنواع أدوات التقييم استخدامًا كانت اختبارات الاختيار من متعدد، يليها اختبارات الصواب والخطأ، ثم اختبارات المزاوجة. وهي بذلك تتفق مع الدراسة الحالية في أن بنود الاختيار من متعدد قد حظيت بأعلى نسبة استخدام. وتختلف عنها في أن بنود الصواب والخطأ والمزاوجة جاءت في المرتبة الثانية والثالثة، بينما حلت في المرتبتين الخامسة والثالثة في الدراسة الحالية.

تختلف الدراسة الحالية مع دراسة الجعافرة السابقة (٢٠٠٩) التي أظهرت نتائجها بأن الأسئلة الموضوعية كانت في معظمها من نوع أسئلة التكميل، بينما جاء الاختيار من متعدد في الدراسة الحالية في المرتبة الأولى، وجاءت التكملة في المرتبة التي تليها.

وتتفق هذه النتائج مع نتيجة الدراسة التي أجراها الدامغ وسكولفيلد (Addamigh & Scholfield 2005) التي توصلت إلى أن أكثر بنود الاختبارات الموضوعية شيوعًا هي بنود الاختيار من متعدد؛ إذ تصدرت القائمة في دراستهما بنسبة قدرها ٣٠,٧٪. أما بقية الأنواع، فقد تشاركت في نسبة ٦٩,٣٪ المتبقية. وجاء نوع تكملة الفراغ في المرتبة الثانية، وبنود المزاوجة في الدرجة الثالثة. وبهذا تتطابق الدراستان في ترتيب الاختبارات الموضوعية وشيوعها كالتالي: الاختيار من متعدد، فالتكملة، ثم المزاوجة.

تتفق الدراسة الحالية مع دراسة هاشم وآخرين (د. ت) في أن أسئلة الاختيار من متعدد أكثر الأسئلة الموضوعية شيوعًا. وتتفق مع دراسة بشير (١٩٩٥) التي تشير إلى أن بنود الاختيار من متعدد هي أكثر الاختبارات الموضوعية فعالية؛ وبأنها تصلح لقياس مستويات التعليم العليا، بالإضافة إلى المستويات الدنيا بالطبع. وتتفق كذلك مع دراسة آيتكين Aitkin (١٩٧٩) التي تشير إلى أن أفضل أساليب الاختبارات الموضوعية يتمثل في الاختيار من متعدد بصوره المختلفة.

أما بنود التكملة، التي حلت في المرتبة الثانية بنسبة عامة بلغت ١٧٪. وهذه النسبة أقل من نصف النسبة التي استأثرت بها بنود الاختيار من متعدد. وهذه النتيجة توضح أهمية بنود التكملة من حيث استخدامها في اختبارات السلاسل على وجه الخصوص. وتتفق هذه النتيجة مع ما توصلت إليه نتائج الدراسة التي قام بها الكحلوت (٢٠٠٢) التي توصلت إلى أن هذه البنود أكثر صعوبة من بنود الاختيار من متعدد، ولكنها أيضًا أكثر تمييزًا. وتتفق هذه النتيجة كذلك مع نتيجة آيتكين Aitkin ، حيث أشارت إلى أن ثاني أنجع الأساليب هو أسلوب التكملة. وفي دراسة الجعافرة السابقة (٢٠٠٩) أظهرت النتائج بأن الأسئلة الموضوعية كانت في معظمها من نوع أسئلة التكميل، بنسبة بلغت ٧٠,٣٪. وبذلك تضع بنود التكملة في الصدارة، وهي تختلف مع الدراسة الحالية التي تضعها في المرتبة الثانية.

استخدمت بنود المزاوجة، التي حلت في المرتبة الثالثة بنسبة عامة بلغت ١٦,٣٪. وهي نسبة قريبة من نسبة استخدامها في دراسة فضل السابقة. وبذلك تتفق هذه النتيجة مع الدراسة التي قام بها الدامغ وسكولفيلد (Addamigh-Scholfield 2005) التي وجدت بأن بنود المزاوجة تأتي في الدرجة الثالثة. وهناك شك في أن تقيس بنود المزاوجة المستويات المعرفية العليا كما أشارت إلى ذلك دراسة بشير المشار إليها في السابق التي توصلت إلى أن هذه البنود غير مناسبة لقياس مستويات التعليم العليا.

تختلف الدراسة الحالية عن دراسة الزبون (٢٠١٣) التي جاءت فيها اختبارات الصواب والخطأ في المرتبة الثانية بعد اختبارات الاختيار من متعدد، بينما حلت في المرتبة الخامسة في هذه الدراسة.

أما إعادة الترتيب الذي جاء في ذيل القائمة، فلم يتناوله أيٌّ من الدراسات السابقة، وترى الدراسة الحالية أنه من الممكن الاستفادة منه في بعض اختبارات العناصر والمهارات؛ مثل: اختبارات الألفاظ (المفردات) والقواعد والتراكيب، واختبارات الكتابة للصفوف المبتدئة، ولكن لا يعول عليها في اختبارات الكتابة الإبداعية.

وللإجابة عن السؤال الثاني: ما الأهداف المعرفية التي تشيع في الاختبارات الموضوعية وتلك التي لا تقيسها؟

ستُعرض الإجابة عن هذا السؤال من خلال النظر في البيانات التي تجمعت في الجداول التالية.

الجدول رقم (٣) ترتيب المستويات المعرفية وَفْقَ الشيوع في السلاسل.

المستويات المعرفية	عدد البنود	النسبة	الرتبة
الفهم	١٧٢٩	٥٨,٩٪	١
التذكر والمعرفة	٩١٧	٣١,٣٪	٢
التطبيق	٢٣٨	٨,١٪	٣
التركيب	٣٧	١,٣٪	٤
التحليل	١٢	.٤٪	٥
التقويم	-	-	-
العدد الكلي للبنود	٢,٩٣٣	١٠٠٪	-

يُستنتج من الجدول رقم (٣) أن قياس مستوى الفهم قد جاء في المرتبة الأولى، بنسبة تتجاوز النصف ٥٨,٩٪. وتلاه في المرتبة الثانية مستوى التذكر بنسبة شيوع قدرها ٣١,٣٪، ثم التطبيق بنسبة ٨,١٪، ومستوى التركيب بنسبة ١,٣٪، وأخيرًا مستوى التحليل بنسبة لم تتجاوز الـ٤.٠٪، وهذه نسبة ضعيفة جدًا لا يعول عليها.

وعليه، نستنتج بأن البنود المستخدمة قد استحوذت عليها عمليات استدعاء المعرفة بصورة عامة، وأغلبها يقيس نواتج التعلم في مدىً يقع بين أدنى سلم النواتج والمتوسط منها؛ مما لا يتوافق مع نواتج التعلم المنشودة.

الجدول رقم (٤) الأهداف المعرفية التي استهدفتها البنود الموضوعية ونسبها.

المستوى	س١	س٢	س٣	س٤	س٥	التكرار ٢,٩٣٣
التذكر والمعرفة	١٣٦ ٪١٤,٨	٨٥ ٪٩,٣	٩٩ ٪١٠,٨	١٨٢ ٪١٩,٨	٤١٥ ٪٤٥,٣	٩١٧ ٪١٠٠
الفهم	٤٤٢ ٪٢٥,٦	٧٠١ ٪٤٠,٥	١٢٨ ٪٧,٤	٣٧٠ ٪٢١,٤	٨٨ ٪٥,١	١٧٢٩ ٪١٠٠
التطبيق	–	–	١٤ ٪٥,٩	٨٠ ٪٣٣,٦	١٤٤ ٪٦٠,٥	٢٣٨ ٪١٠٠
التركيب	–	–	٧ ٪١٨,٩	٣٠ ٪٨١,١	–	٣٧ ٪١٠٠
التحليل	–	–	–	–	١٢ ٪١٠٠	١٢ ٪١٠٠
التقويم	–	–	–	–	–	–

يظهر في الجدول رقم (٤) أن مستوى الفهم، كان المستهدف الأول للبنود الموضوعية؛ فقد استأثر بالنصيب الأوفر من اهتمام معدي هذه الاختبارات بما مقداره ١,٧٢٩ بندًا من المجموع الكلي للبنود البالغ ٢,٩٣٣ بندًا، وكانت أعلى

نسبة لاستخدامه في السلسلة رقم (٢) بما نسبته ٤٠,٥٪، وكان أقل استخدام له في السلسلة رقم (٥) بنسبة لا تتجاوز ٥,١٪.

حل مستوى التذكر والمعرفة في المرتبة الثانية بعدد ٩١٧ بندًا من المجموع الكلي للبنود، وكانت أعلى نسبة لاستخدامه في السلسلة رقم (٥) بنسبة ٤٥,٣٪، وكان أقل استخدام له في السلسلة رقم (٢) بنسبة ٩,٣٪.

جاء مستوى التطبيق تاليًا لمستوى التذكر والمعرفة؛ بعدد ٢٣٨ بندًا، وكانت أعلى نسبة لاستخدامه في السلسلة رقم (٥) بنسبة ٦٠,٥٪، وكان أقل استخدام له في السلسلة رقم (٣) بنسبة تقرب من الـ ٦ ٪. ولم يستخدم قط في السلسلتيْن رقم (١) ورقم (٢).

وأما مستويا التركيب والتحليل، فلم يستهدفا إلا بِعدد قليلٍ جدًّا (٣٧ و١٢) بندًا فقط على التوالي، مما مجموعه ٢,٩٣٣ بندًا، ولم يستخدم مستوى التركيب في ثلاث من السلاسل هي: ١ -٢- ٥، ولم يستخدم التحليل إلا في السلسلة رقم (٥)؛ وقد تبيَّن للدارس أنه، حتى تلك الاثنتيْ عشرة بندًا التي استخدمت في التحليل، في تلك السلسلة، كانت في أدنى مستويات التحليل، فقد كان المطلوب تنفيذه أداءً كتابيًّا آليًّا؛ بتحليل المفردة المكتوبة المنظورة إلى حروفها المنفصلة.

وهكذا، يُستنَج من تحليل البيانات التي يتضمنها الجدول، أن هذه الاختبارات لم تقس قط النواتج عند أرفع المستويات، ألا وهو مستوى التقويم وما يترتب على ذلك من قصور في توظيف مهارات التفكير، والاستدلال، وحل المشكلات.

تنتقل الدراسة الآن إلى عرض مدى شيوع مستويات الأهداف المعرفية في البنود الموضوعية؛ من خلال مدى استخدام كل منها في السلاسل الخمس.

أولًا: الاختيار من متعدد

الجدول رقم (٥) نسب شيوع المستويات المعرفية في بنود الاختيار من متعدد.

السلسلة	س١	س٢	س٣	س٤	س٥	التكرار	النسبة ٪
المستوى المعرفي	٢٢٦ ٪٢٠,٧	٣٢٣ ٪٢٩,٦	٢٠١ ٪١٨,٤	٢٠٧ ٪١٩	١٣٤ ٪١٢,٣	١٠٩١ ٪١٠٠	٪١٠٠
التذكر والمعرفة	٥١	٣٤	٨٠	٥٢	١٠٦	٣٢٣	٪٢٩,٦
مستوى الفهم	١٧٥	٢٨٩	١٠٠	١٢٥	١٥	٧٠٤	٪٦٤,٥
مستوى التطبيق	–	–	١٤	٢٠	١٣	٤٧	٪٤,٣
مستوى التركيب	–	–	٧	١٠	–	١٧	٪١,٦
مستوى التحليل	–	–	–	–	–	–	–
مستوى التقويم	–	–	–	–	–	–	–

يتضح من الجدول رقم (٥) أن بنود الاختيار من متعدد، استأثرت بأعلى نسبة استخدام بعدد ١٠٩١ بندًا من جملة البنود المستخدمة وعددها ٢,٩٣٣ بندًا. ويشيع فيها استخدام مستوى الفهم الذي جاء في المرتبة الأولى بنسبة ٦٤,٥٪، يليها مستوى التذكر والمعرفة بنسبة ٢٩,٦٪، فالتطبيق بنسبة ٤,٣٪، ثم التركيب بنسبة ١,٦٪، لكنها لم تقس قط، المستويْن الأخيريْن (التحليل - التقويم). ويعزى ذلك، في رأي الدارس، إلى أن معدي هذا النوع غالبًا لا تكون لديهم الخبرات والدربة الكافية لتصميم اختبارات تقيس هذه المستويات بهذا النوع من البنود، وليس لأن هذا النوع ليس له القدرة على أن يقيس هذيْن المستويْن. أضف إلى ذلك أن الوقت

الـذي يسـتغرق في تصميمها وبنائها وإعدادها، لا يشـجع على الإقدام على اسـتخدامها في اختبـارات : التشـخيص، والتحصيـل، والتسـكين (تحديـد المسـتوى- التصنيـف) وغير ذلـك مـن الاختبـارات الـتي يعدها مـدرس الصـف على عجلـة مـن أمـره، ولكـن قـد يكـون ذلـك ممكنًـا في الاختبـارات المصيريـة؛ مـن قبيـل اختبـارات الكفايـة المقننـة، والاسـتعداد، والتصفيـة الـتي تجـرى على نطـاق واسـع، ويقـوم على إعدادها فريـق متكامـل مـن خـبراء: الاختبـارات، والقيـاس والتقييـم والتقويـم، واللغـة، وعلـم النفـس والإحصـاء، كمـا هـو الحـال في اختبـارات اللغـة الإنجليزيـة بوصفهـا لغـة أجنبيـة (توفـل TOEFL).

ثانيًا: التكملة

الجدول رقم (٦) نسب شيوع المستويات المعرفية في بنود التكملة.

السلاسل	س١	س٢	س٣	س٤	س٥	التكرار	النسبة ٪
المستوى المعرفي	٥٢ ٪١٠,٤	٧٤ ٪١٤,٨	١٠ ٪٢	١٩٠ ٪٣٨,١	١٧٣ ٪٣٤,٧	٤٩٩ ٪١٠٠	٪١٠٠
التذكر والمعرفة	٧	٥	٦	٥٧	١٢٧	٢٠٢	٪٤٠,٥
مستوى الفهم	٤٥	٦٩	٤	١٠٤	٢٨	٢٥٠	٪٥٠,١
مستوى التطبيق	-	-	-	١٩	١٨	٣٧	٪٧,٤
مستوى التركيب	-	-	-	١٠	-	١٠	٪٢
مستوى التحليل	-	-	-	-	-	-	-
مستوى التقويم	-	-	-	-	-	-	-

يتضح من الجدول رقم (٦) أن بنود التكملة، قد استأثرت بثاني أعلى نسبة استخدام بعدد ٤٩٩ بندًا من جملة البنود المستخدمة. وكما كان عليه الحال في بنود الاختيار من متعدد، فإنه يشيع فيها أيضًا استخدام مستوى الفهم بدرجة عالية، وجاء في المرتبة الأولى بنسبة قدرها ٥٠ ٪ تقريبًا، تلاه مستوى التذكر والمعرفة بنسبة ٤٠,٥٪، فالتطبيق بنسبة ٧,٤٪، وأخيرًا التركيب بنسبة ضئيلة جدًّا، مع ملاحظة أن أربعًا من السلاسل الخمس، لم تستخدم هذا النوع من البنود، كما أن المستويات المعرفية العليا (التحليل والتقويم) لم يتم استخدامهما.

ثالثًا: المزاوجة

الجدول رقم (٧) نسب شيوع المستويات المعرفية في بنود المزاوجة.

السلاسل	س١	س٢	س٣	س٤	س٥	التكرار	النسبة ٪
المستوى المعرفي	٧٧ ٪١٦,١	٢٠٨ ٪٤٣,٥	١٢ ٪٢,٥	١٠٥ ٪٢٢	٧٦ ١٥,٩	٤٧٨ ٪١٠٠	٪١٠٠
التذكر والمعرفة	٤٤	٧	٣	٢٢	٧٦	١٥٢	٪٣١,٨
مستوى الفهم	٣٣	٢٠١	٩	٥٢	–	٢٩٥	٪٦١,٧
مستوى التطبيق	–	–	–	٢١	–	٢١	٪٤,٤
مستوى التركيب	–	–	–	١٠	–	١٠	٪٢,١
مستوى التحليل	–	–	–	–	–	–	–
مستوى التقويم	–	–	–	–	–	–	–

يُستنتج من الجدول أعلاه، أن مستوى الفهم قد استأثر بأعلى نسبة من الاستخدام في بنود المزاوجة بنسبة تجاوزت ٦١,٧٪، وجاء مستوى التذكر في المرتبة الثانية بنسبة ٣١,٨٪، فالتطبيق ٤,٤٪ ثم التركيب ٢,١٪. أما مستويا التحليل والتقويم، فلم يحظيا باهتمام معدي هذه الاختبارات، ويرجح أن يعود السبب في ذلك إلى أن هذا النوع لا يمكن أن تقاس به مهارات التحليل والتقويم التي تحتاج إلى نوع من إعمال الفكر والمقارنة والتحليل وإصدار الحكم، ولا يتأتى هذا إلا من خلال بدائل أخرى غير المزاوجة.

رابعًا: الإجابة القصيرة

الجدول رقم (٨) نسب شيوع المستويات المعرفية في بنود الإجابة القصيرة.

السلاسل	س١	س٢	س٣	س٤	س٥	التكرار	النسبة ٪
المستوى المعرفي	١٠٨ ٢٤,٢٪	١٠٤ ٢٣,٣٪	– –	١٨ ٤٪	٢١٧ ٤٨,٥٪	٤٤٧ ١٠٠٪	١٠٠٪
التذكر والمعرفة	٢٥	٣٩	–	٥	٨١	١٥٠	٣٣,٦٪
مستوى الفهم	٨٣	٦٥	–	١٣	٣١	١٩٢	٤٣٪
مستوى التطبيق	–	–	–	–	١٠٥	١٠٥	٢٣,٤٪
مستوى التركيب	–	–	–	–	–	–	–
مستوى التحليل	–	–	–	–	–	–	–
مستوى التقويم	–	–	–	–	–	–	–

بالنظر في الجدول رقم (٨)، يتضح أن مستوى الفهم مازال يستأثر بالنصيب الأكبر من نسب الاستخدام، وجاء في المرتبة الأولى من حيث متوسط الاستخدام بنسبة قدرها ٤٣٪، تلاه مستوى التذكر بنسبة قدرها ٣٣,٦٪، ثم التطبيق بنسبة ٢٣,٤٪. ومن الجدير بالذكر، أن هذا المستوى لم يُستَخدم إلا في السلسلة رقم (٥). ويلاحظ أن السلسلة الثالثة لم تستخدم فيها بنود الإجابة القصيرة قط. وكما هو الحال مع ما سبق من بنود الاختبارات الموضوعية، فإن المستويات العليا من الأهداف المعرفية (التحليل - التركيب - التقويم) لم تستخدم قط، وذلك يعزى إلى طبيعة هذه البنود واشتراطها بأن تكون الاستجابة كلمة أو عبارة قصيرة محدودة؛ مما يجعل من العسير قياس مهارات التفكير العليا من خلالها؛ لذا على من يحتاج إلى قياس هذه المهارات، اللجوء إلى أدوات قياس بديلة أخرى.

خامسًا: الصواب والخطأ

الجدول رقم (٩) نسب شيوع المستويات المعرفية في بنود الصواب والخطأ.

السلاسل	س١	س٢	س٣	س٤	س٥	التكرار	النسبة ٪
المستوى المعرفي	٦٢ ٪٢٠,٧	٧٧ ٪٢٥,٧	٢٥ ٪٨,٣	١٠٨ ٪٣٦	٢٨ ٪٩,٣	٣٠٠ ٪١٠٠	٪١٠٠
التذكر والمعرفة	-	-	١٠	٣٥	١٤	٥٩	٪١٩,٧
مستوى الفهم	٦٢	٧٧	١٥	٥٠	١٤	٢١٨	٪٧٢,٦
مستوى التطبيق	-	-	-	٢٠	-	٢٠	٪٦,٧
مستوى التركيب	-	-	-	٣	-	٣	٪١
مستوى التحليل	-	-	-	-	-	-	-
مستوى التقويم	-	-	-	-	-	-	-

هذه الطبعة إهـداء من المجمع
ولا يُسمح بنشرها ورقياً أو تداولها تجارياً

يتضح من الجدول رقم (٩) أن مستوى الفهم هو المتصدر لقائمة الاهتمام، واستحوذ على نسبة عالية جدًّا قدرها ٧٢,٦٪، وجاء مستوى التذكر في المرتبة الثانية بنسبة ١٩,٧٪، ثم التطبيق بنسبة ٦,٧٪. وحل مستوى التركيب في المرتبة الأخيرة بنسبة ضئيلة جدًّا قدرها ١٪ ليس إلا. وكما كان الحال في البنود السابقة، فلم يحظ مستويا التحليل والتقويم بأي نصيب من الاستخدام. وقد يعود ذلك إلى طبيعة هذا النوع من البنود الذي يسهل من خلاله قياس المستويات المعرفية الدنيا؛ أيْ مستويَيْ الفهم والتذكر أكثر من قياس مهارات التفكير العليا، التي من الأجدى قياسها ببدائل أخرى قد يكون من بينها اختبارات الكتابة (المقالية) وعمليات التقييم البديل.

سادسًا: إعادة الترتيب

الجدول رقم (١٠) نسب شيوع المستويات المعرفية في بنود إعادة الترتيب.

السلاسل	س١	س٢	س٣	س٤	س٥	التكرار	النسبة٪
المستوى المعرفي	٥٠ ٤٧,٢٪	– –	– –	٣٧ ٣٤,٩٪	١٩ ١٧,٩٪	١٠٦ ١٠٠٪	١٠٠٪
التذكر والمعرفة	٦	–	–	١١	١١	٢٨	٢٦,٥٪
مستوى الفهم	٤٤	–	–	٢٦	–	٧٠	٦٦٪
مستوى التطبيق	–	–	–	–	٨	٨	٧,٥٪
مستوى التركيب	–	–	–	–	–	–	–
مستوى التحليل	–	–	–	–	–	–	–
مستوى التقويم	–	–	–	–	–	–	–

يُستنتج من الجدول رقم (١٠) أن مستوى الفهم قد استحوذ على الصدارة

بنسبة عالية قدرها ٦٦٪، وجاء مستوى التذكر في المرتبة الثانية بعده بفارق كبير بنسبة بلغت ٢٦,٥٪، ثم التطبيق بنسبة ٧,٥٪. ويلاحظ أن السلسلتَيْن الثانية والثالثة لم تستخدما هذه البنود في أيٍّ من المستويات المعرفية. وكما كان الحال في أنواع البنود السابقة، لم تحظ المستويات المعرفية العليا (التحليل - التركيب - التقويم) بأي نصيب من الاستخدام، وقد كان معظم الاستخدام منصبًّا على إعادة ترتيب ألفاظ لتصير جملًا، أو إعادة ترتيب جمل لتصير فقرات.

مناقشة النتائج عن السؤال الثاني

يتضح من مجمل المعطيات الواردة في الجداول السابقة، أن نسبة استخدام المستويات في المجال المعرفي جاءت كالتالي: الفهم بنسبة قدرها ٥٨,٩٪، والتذكر والمعرفة بنسبة ٣١,٣٪، والتطبيق بنسبة٨,١٪، فالتركيب بنسبة ١,٣٪، ثم التحليل بنسبة ٤..٪.

ويلاحظ من مجمل البيانات التي تضمنتها الجداول، أن مستوى الفهم يستحوذ على الصدارة دائمًا، ويحظى بأكثر نسبة من الاستخدام في أنواع البنود الستة دون استثناء. كما يتضح أن مستوى التقويم لم يرد قط في أيٍّ من اختبارات السلاسل قيد الدراسة. وقد يعود ذلك - في المقام الأول - إلى الأهداف التي تنص عليها الخطط الدراسية في المؤسسات التي تتبنى تدريس مثل هذه السلاسل والتي اطلعت الدراسة على الخطط الخاصة بها من خلال ما تتضمنه أدلتها، وقد يعود جزءٌ منه أيضًا، إلى قصور يعُزى إلى قلة الخبرة أو الدربة في تصميم وبناء الاختبارات التي تقيس مستويَيْ التحليل والتقويم؛ مما يجعل الحاجة ماسة لتلافي هذا القصور عند التصدي لإعداد الاختبارات الموضوعية أو في تدريب معديها، أو البحث عن بدائل أخرى تقيس مهارات التفكير العليا.

تؤيد هذه الدراسةُ بعضَ الدراسات السابقة – فيما يخص استخدام البنود الموضوعية في المجال المعرفي – وذلك من حيث اقتصار الاختبارات عمومًا، والاختبارات الموضوعية خصوصًا، على قياس الأهداف في أدنى سلم المجال المعرفي. ففي الدراسة التي أعدها عثمان (٢٠١٥) أوضحت نتائجها أن استخدام مستوى الفهم استأثر بالنصيب الأكبر بنسبة ٤٠٪، تلاه مستوى التذكر بنسبة ٣٠٪، ثم التطبيق بنسبة ٣٠٪ أيضًا، ولم تُستخدم بقية المستويات قط. ومن نقاط الاتفاق الواضحة بين الدراستيْن: (أ) أن مستويات الفهم والتذكر والتطبيق، حلت في المراتب الأولى بالتتالي في كلتا الدراستيْن (ب) وأن نسبة مستوى التذكر في كلتا الدراستيْن كانت قريبة من بعضهما بعضًا؛ فقد جاءت في الدراسة الحالية بنسبة ٣١,٣٪، وفي دراسة عثمان بنسبة ٣٠٪ (جـ) وأن مستوى التقويم لم يستخدم قط في كلتا الدراستيْن.

تتفق الدراسة الحالية مع دراسة أبو عيش (٢٠١٦) التي توصلت إلى أن غالبية مؤشرات أساليب التقويم، تقع بين المتوسط والمرتفع؛ بسبب الاقتصار على أساليب التقويم التقليدية، وأوصت بإقامة دورات في مجال القياس والتقويم تتناول أساليب التقويم الحديثة، وخاصة تلك التي تدعم تنمية التفكير الإبداعي، والتفكير الناقد؛ وهي تشير بذلك – ضمنًا – إلى عدم استخدام المستويات العليا من التحليل والتقويم والإبداع في العينات التي درستها.

وتختلف الدراسة الحالية – جزئيًّا – مع الدراسةِ التي قامت بها دروزة (٢٠١١) التي تشير إلى أن ما توصلت إليه من نتائج، ينسجم والترتيب الهرمي لبلوم؛ حيث كانت أعلى نسبة لمستوى التذكر يليها الفهم، فالتحليل، فالتطبيق، فالتركيب، ثم التقويم، بينما جاء الترتيب هنا كالتالي: الفهم، فالتذكر، فالتطبيق، فالتركيب، ثم التحليل، بينما لم يستخدم مستوى التقويم في الدراسة الحالية قط.

أشارت دراسة الصرايرة (٢٠١١) فيما يتعلق بالنتائج الخاصة بمستويات المجال المعرفي للأهداف، بأن الغالبية العظمى من الأسئلة، قد ركزت على المستويين الأوليْن من مستويات الأهداف المعرفية حسب تصنيف بلوم وهما: التذكر والفهم، في حين كانت الأسئلة التي تقيس أهداف التطبيق متوسطة، وكانت الأسئلة في بقية المستويات العليا قليلة. وبذلك تتفق معها الدراسة الحالية في أن التركيز كان منصبًّا على أدنى مستويَيْن من مستويات الأهداف المعرفية، إلا أنها تختلف معها، في أن مستوى الفهم جاء مقدمًا على التذكر.

واتفقت الدراسة الحالية مع دراسة الجواد (٢٠١١) التي توصلت إلى أن أسئلة الفهم قد احتلت المرتبة الأولى، يليها على الترتيب أسئلة: التذكر، والتركيب، والتحليل، فالتطبيق، ثم التقويم. لكنها اختلفت معها في أن مستوى التقويم لم يستخدم البتة، واختلفت معها أيضًا في أن مستويَيْ التطبيق والتركيب، قد حلّا في المرتبتيْن الثالثة والرابعة، بينما حلّا في دراسة الجواد في المرتبتيْن الثالثة (التركيب) والخامسة (التطبيق). وبينما حل مستوى التحليل في المركز الرابع لديهما، إلا أنه قد حل في المركز الخامس في الدراسة الحالية.

وتتفق هذه الدراسة مع دراسة الجعافرة (٢٠٠٩) في أن أكثر الأسئلة، قد ركزت على المستويين الأوليْن من مستويات الأهداف المعرفية، وهما التذكر والفهم، في حين كانت الأسئلة التي تقيس أهداف التطبيق وسطًا، وكانت في المستويات العليا قليلة. لكنها اختلفت مع ما توصلت إليه هذه الدراسة في أن مستوى الفهم، في الدراسة الحالية، جاء مقدمًا على مستوى التذكر.

كما تتفق مع دراسة حسن (٢٠٠٥) في أن معظم أسئلة الاختبارات تقيس المستويات المعرفية الدنيا، وخاصة مستوى التذكر، دون الاهتمام بالمستويات المعرفية العليا؛ حيث انعدمت النسبة المئوية لكل من مستوى التحليل، ومستوى التركيب، ومستوى التقويم، والاختلاف هنا أن الدراسة الحالية، تقدم فيها مستوى

الفهم على التذكر، وتتفقان في أن مستوى التقويم لم يعالج في كلتا الدراستين.

أشارت دراسة للسليطي (٢٠٠٤) إلى أن الاختبارات الموضوعية تركز على المستويات المعرفية الدنيا، وخاصة مستوى التذكر. وتتفق معها الدراسة الحالية في حقيقة أن الاختبارات تقيس المستويات الدنيا، إلا أن الدراسة الحالية استنتجت أن التركيز في الاختبارات المحللة تركز على مستوى الفهم، وجاء مستوى التذكر في المرتبة الثانية.

كما تؤيد هذه الدراسةُ دراسةً قام بها فضل (٢٠٠٧) استنتجت أن قياس مستوى الفهم قد احتفظ بالرتبة الأولى، سواءً في اختبارات البرامج أو السلاسل أو متوسط الاستخدام، بنسب تجاوزت النصف بقليل، وهو ما توصلت إليه الدراسة الحالية.

استنتجت دراسة الخالدي (١٩٩٥) بأن الأسئلة المضمّنة في الاختبارات التحصيلية، تركز بدرجة عالية على الأسئلة الموضوعية - وعلى وجه الخصوص - أسئلة التذكر التي تتطلب استرجاع المعلومات، والدراسة الحالية تؤيد - جزئيًّا - ما ذهبت إليه تلك الدراسة بفارق أن الدراسة الحالية، تقدم فيها مستوى الفهم على التذكر.

ومن الأعمال التي تتفق معها الدراسة الحالية، بحث مجلة رسالة الخليج العربي (١٩٩٥) الذي توصل إلى أن الاختبارات تقيس المستويات الدنيا في المستوى المعرفي عند تقويم الطلبة. وكذلك دراسة بشير (١٩٩٥) فيما يتعلق بالمستويات المعرفية، حيث أشارت إلى أن بنود الصواب والخطأ، تصلح لقياس المستويات العقلية الدنيا، وهي ليست مناسبة لقياس المستويات العليا. وقد اتفقت كذلك مع الدراسة المسحية التي قام بها القرشي (١٩٨٦) ووجد أن الأساليب المتبعة في القياس والتقويم، تقتصر على الجانب المعرفي غالبًا وفي مستوياته الدنيا، وعزت ذلك إلى عدم ارتباط التقويم بأهداف محددة.

كما تتفق مع دراسة أعدها حطاب (٢٠٠٢) حيث استنتجت أن أسئلة الاختبارات تركزت على المستويات المعرفية الدنيا من تصنيف بلوم للمجال المعرفي، وبخاصة مستوى التذكر، وكان الإهمال واضحًا للمستويات العليا، وبخاصة مستويَيْ التركيب والتقويم. وبذا تتفق معها الدراسة الحالية بأن مستويَيْ التقويم والتحليل لم يحظيا بالاهتمام قط.

عرض النتائج عن السؤال الثالث ومناقشتها

للإجابة عن السؤال الثالث: ما النواتج التعليمية التي ينبغي أن تحققها الاختبارات الموضوعية والعيوب الكامنة فيها؟

ستُعرض الإجابة عن هذا السؤال في ضوء البيانات والمعلومات التي تجمعت في الجداول المستقاة من السلاسل قيد الدراسة، بالإضافة إلى ما ورد في الأدبيات والدراسات التي تناولت الاختبارات المستخدمة.

يتضح من مجمل البيانات التي تتضمنها جداول شيوع المستويات المعرفية في تغطية نواتج التعلم، أن البنود المستخدمة قد استحوذت عليها عمليات استدعاء المعرفة بصورة عامة، وكان أغلب أنواع البنود يقيس نواتج التعلم في مدىً يقع بين أدنى سلم النواتج وقليل من المتوسط منها، وأنها لم تقسْ قط أعلى المستويات (مستوى التقويم) مما يعني عدم توظيف مهارات كالتفكير، والاستدلال، وحل المشكلات؛ وهذا ما لا يتوافق مع نواتج التعلم، ولا يدعم تنمية التفكير الإبداعي.

وباطلاع الدارس على خطط البرامج التي تتبنى السلاسل، وجد أن تحديد نواتج التعلم يتسم بالضعف توصيفًا وتحديدًا، وأن اختيار أساليب القياس والتقييم يترك غالبًا لرؤية المعلم. وأن إعداد الاختبارات لا يقوم على معايير علمية واضحة. وفيما يلي تُستَعرضُ بعض أوجه الاتفاق والاختلاف مع الدراسات السابقة فيما يخص هذه النواتج.

تتفق الدراسة الحالية مع دراسة هاشم وآخرين (د.ت) السابقة، بأن «أهم نقاط الضعف الخاصة بطرق التقويم، أن الامتحانات لا تغطي محتويات المقرر... ويغلب عليها استدعاء المعرفة، وأن من العناصر التي ينبغي مراعاتها في عمليات التقويم، اتساق أسلوب التقويم مع نواتج التعلم» (ص.٤).

تتفق أيضًا مع دراسة أبو عيش (٢٠١٦) التي وجدت أن غالبية مؤشرات أساليب تقويم نواتج التعلم، تقع بين المتوسط والمرتفع، وتختلف عن الدراسة الحالية في أن الأخيرة وجدت أن المؤشرات لم تلامس قط أرفع المستويات، ألا وهو مستوى التقويم وما يترتب على ذلك من عدم توظيف مهارات: التفكير، والاستدلال، وحل المشكلات. ولكنها توافقها في «ضرورة إعداد دورات تنمية مهنية لأعضاء هيئة التدريس في مجال القياس والتقويم بحيث تتناول أساليب التقويم الحديثة، وممارسات تقويمية يمكن استخدامها في تقويم نواتج التعلم، وخاصة تلك التي تدعم تنمية التفكير الإبداعي، والتفكير الناقد» (ص.٦٢).

تتفق الدراسة الحالية مع دراسة أعدها آل داود وآخرون (٢٠١٦)، في أن تحديد نواتج التعلم، للبرامج الأكاديمية والمقررات الدراسية، يؤدي إلى اختيار أساليب التقويم الحديثة للتحقق من اكتساب المتعلمين لنواتج التعلم المقصودة. وتتفق معها أيضًا في أن إجراءات تقويم نواتج التعلم التي بحثوا فيها «لا تقوم على معايير علمية واضحة، وإنما تستند على خيارات ورؤى الأساتذة، دون العمل وفق معايير وأطر موحدة، كما أن التقويم لا يقيس ما ينبغي تحصيله من مهارات ومعارف بحسب أدلة الجامعات والكليات التي اطلعوا عليها» (ص.٥٠٦).

استنتجت دراسة عثمان (٢٠١٥) عن واقع أساليب تقويم نواتج التعلم، أن الاختبارات التي تم فحصها لم تُستخدمْ فيها مستويات: التحليل والتقويم والإبداع. وهذا ما توصلت إليه الدراسة الحالية بأن البنود المستخدمة قد استحوذت عليها عمليات استدعاء المعرفة بصورة عامة، وأغلب أنواع البنود

تقيس نواتج التعلم في مدىً يقع بين أدنى سلم النواتج والمتوسط منها؛ وهذا مما لا يتوافق مع نواتج التعلم المنشودة.

هناك أمر جدير بالذكر والاهتمام، ألا وهو تدريب الأفراد الذين توكل إليهم مهمة تصميم الاختبارات وبنائها وإعدادها. فقد وجدت الدراسة الحالية، كما وجدت دراسات كثيرة سابقة، أن هناك قصورًا فيما يتعلق بالتدريب والتأهيل لكل من يُقْدِمُ على استخدام الاختبارات الموضوعية في القياس والتقييم.

ومن هذه الدراسات دراسة بشير (١٩٩٥) التي تذكر أنه توجد صعوبة في بناء بنود الاختيار من متعدد؛ لأنها تتطلب مهارات لا تتوافر إلا بعد إعداد وتدريب خاصين لواضعيها. كما قدمت دراسة حامد (١٩٩٧) توصيات من أبرزها، ضرورة تدريب المعلمين وتأهيلهم في أساليب التقويم المتنوعة، ولا سيما الاختبارات المقننة. ومنها كذلك دراسة الخالدي (١٩٩٨) التي رأت ضرورة تنظيم دورات تدريبية للمعلمين تركز على كيفية تصميم وبناء الأسئلة المقالية والموضوعية وكيفية تصحيحها وتحليلها وفق الأسس العلمية. كما دعت دراسة نصر (١٩٩٨) القائمين على تعليم اللغة العربية، بضرورة عقد دورات وورش عمل متخصصة لمعلمي اللغة العربية، تتناول أساليب التقويم وأدواته التي يتدنى استخدام المعلمين لها، وإتاحة الفرصة لهم بالتدرب العملي على إعداد نماذج من أدوات التقويم المستهدفة وتطبيقها.

ومن الدراسات التي نادت بضرورة عقد ورش تدريبية أيضًا، دراسة عبد الرازق (٢٠٠٢) الذي قام بدراسة تجريبية درب فيها المعلمين على أساليب الاختبارات المختلفة من خلال ورشة تعليمية أدت إلى تحسن مهاراتهم في صياغة الأسئلة التحريرية التي تقيس مختلف المستويات المعرفية والإدراكية. كما أوصت دراسة آل داود وآخرين (٢٠١٦) بعقد ورش تعليمية لكيفية صياغة نواتج التعلم، وإقامة برامج تدريبية لأعضاء هيئة التدريس في الاتجاهات الحديثة لأساليب التقويم

هذه الطبعة إهـداء من المجمع
ولا يُسمح بنشرها ورقياً أو تداولها تجارياً

ومهارات بنائها.

التوصيات:

في ضوء النتائج التي توصلت إليها الدراسة ومن خلال الإجابة عن أسئلة الدراسة المتعلقة بـ: شيوع البنود، والمستويات المعرفية، ونواتج التعلم، فإنها توصي بما يلي:

١. توجيه القائمين على أمر الاختبارات عمومًا، والاختبارات الموضوعية على وجه الخصوص، بمراعاة التنوع في تضمين كل المستويات المعرفية في أساليب التقييم.

٢. ألّا يكون التركيز أو الاقتصار على نوع واحد من أنواع الاختبارات الموضوعية.

٣. إعداد الاختبارات الموضوعية وَفْقَ معايير محددة، وفي ضوء جداولَ للمواصفات.

٤. ضرورة تزويد القائمين على أمر الاختبارات عمومًا، والاختبارات الموضوعية على وجه الخصوص، بدليل المستويات (بلوم مثلًا) وبدليل النواتج التعلمية التعليمية، ودعوتهم إلى الاهتمام بالتركيز على المستويات العليا لتنمية التفكير الناقد.

٥. إقامة ورش عمل ودورات تدريبية؛ لتدريب القائمين على أمر الاختبارات عمومًا، والاختبارات الموضوعية على وجه الخصوص، على قضايا التنوع والتوازن في إدماج المستويات المعرفية العليا.

٦. تدريب المعلمين على أساليب التقييم التي تقيس نواتج التعلم والتعليم التي يُنَصُّ عليها في الخطط الدراسية.

المراجع

- آل داود، إبراهيم بن محمد بن إبراهيم، والمشاري، أمل إبراهيم بن عامر، وعبد الحميد، أمل عبد المنعم (٢٠١٦). مدى توافق أساليب تقويم نواتج تعلم طالبات كليات العلوم والدراسات الإنسانية بجامعة سطام بن عبد العزيز مع معايير الجودة (تصور مقترح). مجلة التربية. جامعة الأزهر. (١٧١). ديسمبر. الجزء الأول. ٥٠٢ – ٥٣٨.

- أبو علام، رجاء محمد. (١٩٨٧). قياس وتقويم التحصيل الدراسي. دار القلم.

- أبو عيش، بسينة رشاد بن علي. (٢٠١٦). أساليب تقويم نواتج التعلم لدى طلبة جامعة الطائف في ضوء معايير الجودة. مجلة العلوم التربوية. (٢٧). ١٣-٦٧. جامعة جنوب الوادي. كلية التربية.

- بركات، محمد خليفة. (١٩٩٤) مناهج البحث العلمي في التربية وعلم النفس. دار القلم.

- الجعافرة، خضراء أرشود. (٢٠٠٩). دراسة تحليلية لأسئلة كتب اللغة العربية للصفوف الخامس والسادس والسابع في الأردن. مجلة العلوم التربوية والنفسية. ١٠ (٤) ٦٣-٨٦.

- الجواد، إياد، وقنديل، أنيسة. (٢٠١١). دراسة تقويمية للاختبارات التحصيلية في اللغة العربية للشهادة الثانوية بفلسطين ومقترحات لتطويرها. مجلة القراءة والمعرفة. الجمعية المصرية للقراءة والمعرفة، كلية التربية. جامعة عين شمس، العدد (١١١) الجزء الأول، ٢٣٣- ٢٥٩.

- حطاب، مهدي صحي، والفقي، عبد المؤمن فرج. (٢٠٠٢). تقويم أساليب قياس الجانب المعرفي في المعهد العالي لإعداد المعلمين في بنغازي. المجلة العربية للتربية. ٢٢ (١). ٥١-٥٨.

- الخالدي، ابتسام عبد الأمير، والسويدي، موزه. (١٩٩٨). دراسة مدى فاعلية أساليب القياس والتقويم السائدة في مدارس دولة الإمارات العربية المتحدة. قسم البحوث التربوية، وزارة التربية والتعليم والشباب.

- الدامغ، خالد بن عبد العزيز، وفضل، محمد. (٢٠١٢). الميسر في إعداد الاختبارات لمدرسي اللغات الأجنبية.

- دروزة، أفنان نظير. (٢٠١١). درجة مراعاة المعلمين في مدارس قلقيلية لمستويات بلوم للأهداف المعرفية لدى تخطيطهم للتدريس. مجلة جامعة النجاح للأبحاث والعلوم الإنسانية. ٢٥(١٠) ٢٥٥٩-٢٥٨٢. جامعة النجاح.

- الزبون، حابس سعد. (٢٠١٣). مدى كفاءة معلمي المدارس الحكومية الثانوية التابعة لإدارة التربية والتعليم بالطائف في بناء الاختبارات التحصيلية وفق معايير الاختبار الجيد. مجلة كلية التربية، (١٤). جامعة بورسعيد.

- السديس، عبد الله محمد. (٢٠٠١). أساليب التقويم المتبعة في مهارة القراءة بمعاهد اللغة العربية - جامعة الإمام محمد بن سعود الإسلامية، [رسالة ماجستير غير منشورة].

- السليطي، حمدة حسن، وتايه، خضر عبد الله. (٢٠٠٤). دراسة تحليلية تقويمية لأسئلة اختبارات الشهادة الثانوية العامة لمقررات اللغة العربية. مجلة رسالة الخليج العربي، (٩٣). ٩٧-١٣٣.

- سمارة، عزيز، والنمر، عصام، وإبراهيم، محمد عبد القادر. (٢٠٠٢) مبادئ القياس والتقويم في التربية (ط٢). دار الفكر.

- الصرايرة، آيات جعفر. (٢٠١١). دراسة تحليلية لأنماط أسئلة الاختبارات التحصيلية لدى معلمي التربية الاجتماعية والوطنية للصفين الرابع والخامس الأساسيْن في لواء المزار الجنوبي [رسالة ماجستير غير منشورة]. كلية الدراسات العليا. جامعة مؤتة.

- طعيمة، رشدي أحمد. (٢٠٠٣). نماذج من الاختبارات الموضوعية في اللغة العربية. دار الفكر العربي.
- عبد الرازق، صلاح عبد السميع. (٢٠٠٠). تنمية مهارات صياغة الأسئلة التحريرية. دار القاهرة.
- عبده، مصطفى. (٢٠٢١). التعليم المبني على نواتج التعلم.
https://www.teachingacademy2021.com
- عثمان، يسري محمد. (٢٠١٥، ديسمبر). واقع أساليب تقويم نواتج التعلم لدى طلبة السنة التحضيرية بجامعة الدمام وفق متطلبات الجودة والاعتماد الأكاديمي [عرض ورقة عمل]، المؤتمر الثاني للقياس والتقويم. الرياض.
- علام، صلاح الدين محمود. (٢٠٠٤). التقويم التربوي البديل: أسسه النظرية والمنهجية وتطبيقاته المنهجية. دار الفكر العربي.
- عودة، أحمد سليمان. (١٩٨٩). دراسة مقارنة أثر اختبار الاختيار من متعدد واختبارات الإجابة القصيرة في الاحتفاظ بنواتج التعليم في المجال المعرفي. المجلة التربوية، جامعة الكويت، (١٩). ١٠٩-١٢٧.
- فضل، محمد عبد الخالق. (٢٠٠٧). الاختبارات الموضوعية في برامج تعليم اللغات الأجنبية [أطروحة دكتوراه غير منشورة]. كلية التربية، جامعة أفريقيا العالمية.
- فضل، محمد عبد الخالق. (٢٠٠٧). أوجه القصور في الاختبارات الموضوعية. مجلة اللغة العربية لغير الناطقين بها، معهد اللغة العربية، جامعة أفريقيا العالمية.

- فضل، محمد عبد الخالق. (٢٠١٦). واقع اختبارات اللغة العربية في العالم العربي، في « اختبارات اللغة: تجارب وآفاق». مركز الملك عبد الله الدولي لخدمة اللغة العربية. ١١-٥٥.
- الكحلوت، أحمد إسماعيل (٢٠٠٢). مقارنة بين الخصائص السيكومترية لكل من اختبارات الاختيار من متعدد واختبارات التكميل. مجلة مركز البحوث التربوية، جامعة قطر. (٢٢). ١٢٧-١٥٦.
- المركز العربي للبحوث التربوية بدول الخليج. (١٩٩٥). تطوير أساليب تقويم الطلبة بمراحل التعليم العام. مجلة رسالة الخليج العربي. (٩٣). ٩٧-١٣٣.
- هاشم، إكرام علي، وسيد، أحمد حسن، وكمال، مصطفى محمد. دليل تقويم نواتج التعلم، مشروع تطوير نظم تقويم الطلاب والامتحانات. كلية التمريض. جامعة أسيوط. https://life.aun.edu.eg
- Addamigh, Khalid & Phil Scholfield. (2005). Test Item format in EFL Tests at Secondary Schools and University Levels in Saudi Arabia". TESOL Arabia. 91-111.
- Aitkin, K.)1979(. Techniques of Assessing Listening Comprehension in Second Languages. Audio-Visual Language Journal, 17 (3) 119-125.
- Gronlund, N. (1990). Measurement and Evaluation in Teaching (7th ed). New Macmillan Publishing Co.
- Osterlind, Steven. (1992). Constructing Test Items (2nd ed.). Publishers, Norwell, MA.24
- Stiggins, Student. (2001). Involved Classroom Assessment (3rd ed.). Upper Saddle River.

تصور مقترح لبناء اختبار معياري لقياس كفايات معلمي اللغة العربية لغة ثانية

د:عبيد الله بن عبدالله الفايدي[1]

مستخلص

يهدف هذا البحث إلى تقديم تصور مقترح لبناء اختبار معياري لقياس كفايات معلمي اللغة العربية لغة ثانية، وإعداد قائمة بالكفايات المهنية واللغوية والثقافية اللازمة لمعلم اللغة العربية لغة ثانية، وللوصول إلى نتائج البحث استخدم الباحث المنهج الوصفي التحليلي؛ لإعداد قائمة بالكفايات اللازمة لمعلمي اللغة العربية لغة ثانية، وتوصل البحث إلى قائمة بعدد من الكفايات تمثلت في:

- الكفايات المهنية:

وانبثق عنها (٤) كفايات فرعية، تضمنت (٦٤) كفاية إجرائية.

- الكفايات اللغوية:

وانبثق عنها: (٤) كفايات فرعية تضمنت (٦٤) كفاية إجرائية.

- الكفايات الثقافية:

وانبثق عنها (٤) كفايات فرعية، تضمنت (٥٢) كفاية إجرائية.

فمجموع الكفايات الإجرائية التي توصل إليها الباحث: (١٨٠) كفاية، وقدَّم تصورا مقترحا لبناء اختبار لقياس هذه الكفايات، وبناء على النتائج التي توصل إليها قدَّم عددًا من التوصيات والمقترحات.

---

(١) أستاذ مساعد بمعهد تعليم اللغة التابع للجامعة الإسلامية في المدينة المنورة

اختبارات اللغة
Journal of Language Testing

الكلمات المفتاحية:

كفايات معلمي اللغة العربية، الاختبارات المعيارية، الكفاية اللغوية، الكفاية الثقافية، الكفاية المهنية.

Abstract:

The aim of this research is to present a proposal for building a standard test to measure the competencies of teachers of Arabic as a second language, and to prepare a list of the professional, linguistic and cultural competencies necessary for these teachers. To reach the results of this research, the researcher used the descriptive analytical approach for preparing a list of the necessary competencies and reaching a list of a number of competencies, they are:

1. Professional competencies from which (4) sub-competencies emanated, from these sub- competencies emanated (64) procedural competencies.
2. Linguistic competencies from which (4) sub-competencies emanated, and (64) procedural competencies.
3. Cultural competencies from which (4) sub-competencies emanated, and (52) procedural competencies.

The total of procedural competencies are (180). The researcher presented a proposed concept to build a test to measure these competencies, and based on the results he reached, he presented a number of recommendations and proposals.

Keywords:

competencies of teachers of the Arabic language, standardized tests, linguistic competence, cultural competence, professional competence.

مقدمة:

المعلم هو العنصر الرئيس الذي تقوم عليه العملية التعليمية في كافة الجهات التعليمية؛ لذا فإن تأهيله المهني واللغوي والثقافي ضرورة ملحة، إضافة إلى أنَّ الجهات التعليمية التي تحرص على نواتج تعلم جيدة تسعى لاستقطاب المعلمين الأكفاء الذين يسهمون في تحقيق نواتج التعلم المستهدفة لتلك الجهات؛ وهذا مما يؤكد على ضرورة أن يكون اختيار المعلم مبنيًّا على أسس علمية.

ويتوقف التنفيذ الفعلي في قاعات الدرس عليه، فإعداد المعلم لاسيما في تعليم اللغة -أي لغة- يحتاج إلى إعداد علمي خاص، في اللغة المستهدفة؛ مهارتها، وعناصرها، وفي علم اللغة التطبيقي خاصة، مع ضرورة تدريبه أثناء العمل؛ ليطلع على كل جديد في المجال، ولتنمية خبراته التدريسية ومهاراته. (الراجحي، ١٩٩٠: ١٢٤).

وهو القدوة والنموذج وصانع التغيير، وبقدر ما يملكه من فَهْمٍ سليمٍ لقواعد المهنة، وأساليب ممارستها، وضرورة تطويرها، والتطور معها، والالتزام بأدائها على أحسن صورة، بقدر ما يكون قادرًا على النجاح، وتحقيق الآمال المعقودة عليه، وتخطي التحديات التي تقف في طريق التميز والإبداع والمنافسة في ميادين المهنة؛ وهذا ما يحتّم على كل من يعمل في مهنة التعليم أن يتمثَّل المقومات العلمية والأخلاقية والفكرية التي تجعل منه قوة دافعة نحو النهوض وتعزيز فرص النجاح. (هيئة تقويم التعليم والتدريب، ٢٠١٧: ١١)

علاوة على أنه يؤدي دورًا مهمًا في منظومة التدريس؛ إذ إنه منوط به أن يحقق نواتج التعلم المستهدفة في جوانبها المختلفة (المعرفية، المهارية، الوجدانية) لدى المتعلمين، خاصة في هذا العصر الذي تعددت فيه الروافد التي يمكن أن تسهم بصورة مباشرة أو غير مباشرة في تكوين شخصية المتعلم؛ لذا على المعلم أن يستثمر ويوظف تلك الروافد في التعليم (الحديبي، ٢٠٢٠: ١٣)

اختبارات اللغة
Journal of Language Testing

وفي عصر يتَّسِم بسرعة التغيير والتنافسية العالية، والتَّوجه نحو الثورة الصناعية الرابعة، والإفادة منها ومما أفرزته من علوم ومناهج علمية تجريبية، تسعى المملكة العربية السعودية إلى بناء وطن مزدهر يقف في مقدمة دول العالم، ويكون له نصيبٌ وافرٌ من التقدم وتحقيق الريادة والتَّميز وهي نقلةٌ تضمن -بعد توفيق الله- تحولًا إلى تعليم متقدم وأكثر فاعلية في شتى المجالات، يلبّي متطلبات القرن الحادي والعشرين، ويواكب التطورات العالمية والثورة المعرفية والتقنية؛ فالتعليم هو السبيل لبناء جيل واعٍ قادر على تمثُّل ثقافته أولًا، وعلى التكيف والتفاعل مع الثقافات الأخرى وما تتضمنه من معطيات ثانيًا. (هيئة التعليم والتقويم، ٢٠١٧: ١١)

ولأن مهنة التعليم تستهدف تكوين شخصية المتعلم قبل التخصص في أي مهنة؛ فإن أي إصلاح للتعليم أو تطوير له إنما ينطلق من البصمات التي يتركها المعلم في معارف وسلوكيات المتعلمين ومهاراتهم، وقِيَمهم، واتجاهاتهم، وطرائق تفكيرهم واستدلالهم، وبناء أحكامهم؛ تماشيًا مع كون الاهتمام بمهنة التعليم في أي مجتمع يشير إلى مدى تحمله مسؤوليته تجاه بناء أجياله. (هيئة تقويم التعليم والتدريب، ٢٠١٧: ١١)

وهذا يتطلب من جهات ترشيح المعلمين وتعيينهم أن تحدد عددًا من الكفايات اللازمة، التي يجب على المعلم أن يكون قادرًا على توظيفها في العملية التعليمية؛ ليكون أداؤه فاعلًا، ويمكن قياس الكفايات من خلال اختبار معياري معتمد ومصمم لقياسها؛ فقد يسهم ذلك في تحقيق الأهداف المنشودة، وينمي مهارات المتعلمين اللغوية والتواصلية وفق أحدث الاتجاهات، ويواكب التطورات التعليمية الحديثة، والذي لا شك فيه أن المعلم الذي تتوافر فيه هذه الكفايات المهنية واللغوية والثقافية أجدر للقيام بهذه المهمة.

من هنا فإنَّ معلم اللغة العربية لغة ثانية يجب أن تتوفر فيه عدد من الكفايات المهنية واللغوية والثقافية التي تؤهله لأن يكون معلمًا فاعلًا في تدريسه،

وهذا يتطلب إعدادًا تربويًّا ولغويًّا وثقافيًّا من خلال الجهات المؤهلة للمعلمين في مجال تعليم اللغات، والإلمام بهذه الكفايات مطلب أساسي في معلم اللغة العربية لغة ثانية كي يكون تعليمها جيدًا وفاعلًا.

فالاختبارات المهنية المبنية على أسس علمية دقيقة تُعدّ وسيلة لقياس تمكن المعلم من الأداء المناسب في مجال تعليم اللغة العربية لغة ثانية، فقد لاحظ الباحث من خلال تدريبه لعدد من المعلمين من جنسيات مختلفة أن بعضهم قد رشح لتعليم اللغة العربية مع نقص في خبراته المهنية والثقافية واللغوية، وهذا لا شك مما يجعل تعلُّم اللغة وتعليمها أطول وقتًا، وأكثر جهدًا، وأقل جودة، إلا فيما ندر.

لذا يرى الباحث أن بناء اختبار معياري لقياس كفايات معلمي اللغة العربية لغة ثانية، ومدى تمكنهم من توظيف الكفايات المهنية واللغوية والثقافية، قد يحقق الأهداف المنشودة في نشر اللغة العربية وإبراز مكانتها، من خلال ترشيح أكفاء لتدريسها وفق أحدث الأساليب والطرق، وبما يتناسب مع التوجهات الثقافية المحلية والعالمية، التي تسهم في بناء شخصية المتعلم وتحقيق أهدافه من تعلم اللغة.

مشكلة البحث:

دفع الباحث لإجراء هذا البحث وتقديم هذا المقترح لبناء اختبار معياري لقياس الكفايات اللازمة لمعلمي اللغة العربية لغة ثانية، عدة أمور؛ منها:

١. **ما أشارت إليه بعض الدراسات من أهمية تأهيل معلمي اللغة العربية لغة ثانية:**

حيث توصلت دراسة (حامد؛ آخرون، ٢٠١٤)، إلى: ضرورة تطبيق نظام الترخيص لممارسة مهنة التعليم لما يمثله من نظام مقنن يخضع المتقدمين لعمليات فرز وانتقاء دقيقة تؤدي إلى تجويد العملية التعليمية بأكملها.

ودراسة (الحديبي، ٢٠١٧) التي تناولت المشكلات التي تواجه الجهات المعنية بتعليم اللغة العربية للناطقين بلغات أخرى، وطبقت على عينة من المختصين بلغ عددهم (١١٥) مختصًا في مجال تعليم اللغة العربية للناطقين بلغات أخرى، وتوصلت هذه الدراسة إلى أن هناك (٢٥) مشكلة تتعلق بالمعلم من حيث إعداده وتدريبه، من هذه المشكلات: عدم وجود معلمين حاصلين على شهادة تخصصية في (تعليم اللغة العربية للناطقين بغيرها).

ودراسة (الحديبي، ٢٠٢٠) التي اقترحت تقويم مهارات معلمي اللغة العربية على ضوء المعايير المعاصرة لإعداد معلم اللغة العربية لغة ثانية.

٢. **ندرة الدراسات التي تناولت بناء اختبار معياري لقياس كفايات معلمي اللغة العربية لغة ثانية، حيث لم يعثر الباحث على أي دراسة تناولت بناء اختبار معياري لقياس كفايات معلمي اللغة العربية لغة ثانية.**

٣. **حاجة المجال لإجراء مثل هذا الاختبار ليكون معيارًا ومرجعية علمية معتمدة لتعيين معلمي اللغة العربية لغة ثانية.**

من هنا سعى الباحث لبناء قائمة بالكفايات اللازمة لمعلمي اللغة العربية لغة ثانية، ليتم على أساسها بناء اختبار معياري لقياس الكفايات المهنية واللغوية والثقافية اللازمة للمتقدمين للعمل في مجال تعليم اللغة العربية لغة ثانية؛ ليتم الاختيار على أسس علمية دقيقة، وقدم تصورًا مقترحًا لبناء هذا الاختبار.

وتحددت مشكلة البحث الحالي في عدم وجود اختبار معياري -على حد علم الباحث- لقياس الكفايات اللازمة لمعلمي اللغة العربية لغة ثانية، وللتصدي لهذه المشكلة سعى الباحث للإجابة عن السؤالين الآتيين:

٤. **ما الكفايات اللازمة التي ينبغي توفرها في معلمي اللغة العربية لغة ثانية؟**

٥. **ما التصور المقترح لبناء اختبار معياري لقياس الكفايات اللازمة لمعلمي اللغة العربية لغة ثانية؟**

أهداف البحث:

يهدف البحث الحالي إلى:

١. إعداد قائمة بالكفايات اللازمة التي ينبغي توافرها في معلمي اللغة العربية لغة ثانية.

٢. تقديم تصور مقترح لبناء اختبار معياري لقياس كفايات معلمي اللغة العربية لغة ثانية.

حدود البحث:

التزم البحث الحالي بالحدود الآتية:

١. **الحدود الموضوعية:**

- قائمة الكفايات اللازمة لمعلمي اللغة العربية لغة ثانية.
- بناء تصور مقترح لبناء اختبار معياري لقياس كفايات معلمي اللغة العربية لغة ثانية.

٢. **الحدود البشرية:** معلمو اللغة العربية لغة ثانية، والراغبون في تعليم اللغة العربية لغة ثانية.

٣. **الحدود الزمنية:** الفصل الدراسي الثاني لعام ١٤٤٤هـ.

أهمية البحث:

تنبع أهمية هذا البحث من أهمية الكفايات المهنية واللغوية والثقافية اللازمة لمعلمي اللغة العربية لغة ثانية.

إضافة إلى أنه قد يفيد:

- **في التخطيط الجيد المبني على أسس علمية في جانب الاكتساب اللغوي ونشر اللغة العربية في أنحاء العالم** من خلال ترشيح مختصين مؤهلين مهنيًا ولغويًا وثقافيًا يُعلمون اللغة العربية لغة ثانية بما يتناسب مع التوجهات الحديثة في تعليم اللغات لغة ثانية.
- **معاهد تعليم اللغة العربية لغة ثانية داخل وخارج المملكة، ومراكزها؛** إذ يفيد هذا التصور في تقديم إجراءات بناء اختبار معياري لقياس الكفايات اللازمة لمعلمي اللغة العربية لغة ثانية؛ ليكون الترشيح في تعليمها مبنيًا على أسس علمية تحقق الموضوعية في الاختيار.
- **في تطوير برامج تعليم اللغة العربية لغة ثانية:** من خلال تضمين الكفايات الإجرائية التي توصل إليها البحث في برامج الإعداد المهني واللغوي والثقافي للراغبين في الالتحاق بالبرامج التأهيلية في إعداد معلمي اللغة العربية لغة ثانية.
- **مصممي الاختبارات المعيارية ومطوريها في مجال تعليم اللغة العربية لغة ثانية،** من خلال قائمة الكفايات الإجرائية اللازمة لمعلمي اللغة العربية لغة ثانية.
- **عمل مشروع لبناء اختبار معياري لقياس الكفايات اللازمة لمعلمي اللغة العربية لغة ثانية.**

- **القائمين على برامج تدريب معلمي اللغة العربية لغة ثانية**، وذلك من خلال إجراء اختبار معتمد لتحديد نقاط القوة ونقاط الضعف، وبناء على نتائج المعلمين في الاختبار يتم تقديم البرامج التدريبية التي تلبي احتياجاتهم.

مصطلحات البحث:

- **الكفاية:** مجموعة المعارف والاتجاهات والمهارات والخبرات التي يمتلكها معلمو اللغة العربية لغة ثانية، ويوظفونها في أداء مهماتهم التعليمية بدرجة عالية من الإتقان؛ من أجل تحقيق الأهداف التعليمية المنشودة. (الربابعة، ٢٠١٦: ١٦٥٤)

- **كفايات معلمي اللغة العربية لغة ثانية:** مجموعة المعارف والخبرات والأداءات المعرفية والعملية التي يوظفها المعلم أثناء تعليم اللغة العربية لغة ثانية، وهذه الكفايات: مهنية، ولغوية، وثقافية، يتضمن كل منها كفايات فرعية وكفايات إجرائية أدائية يؤديها المعلم داخل حجرة الصف، لتنمية مهارات متعلمي اللغة العربية لغة ثانية وفق أحدث التوجهات في مجال تعليم اللغات.

- **الاختبار المعياري يعرفه الباحث إجرائيًّا بأنه:** مجموعة من الأسئلة الاختبارية التي تستهدف قياس كفايات معلمي اللغة العربية لغة ثانية؛ ليكون مرجعية علمية معتمدة في ترشيح معلمي اللغة العربية لغة ثانية، ويجب أن يتضمن هذا الاختبار الكفايات الإجرائية التي ينبغي أن يراعيها معلم اللغة العربية لغة ثانية أثناء تعليمها.

اختبارات اللغة
Journal of Language Testing

الإطار النظري للبحث، والدراسات السابقة:

أولا: الإطار النظري للبحث:

المحور الأول: مفهوم كفايات معلمي اللغة العربية لغة ثانية:

لتحقيق الأهداف التعليمية والتدريسية يجب على الجهات التعليمية أن تركز على المعلم وتطويره وحسن اختياره للقيام بالمهام التدريسية كما ينبغي، إضافة إلى أنَّ تطوير المعلم، وتنمية معارفه وصقل مهاراته، ضرورة ملحة؛ ليتمكن من أداء رسالته التعليمية بما يتلاءم مع متطلبات العصر واحتياجات المتعلمين؛ لذا ينبغي أن يكون اختيار المعلم بناء على أسس علمية في الاختيار والتعيين ولا تقتصر على التخصص في اللغة العربية، لاسيما في مجال تعليم اللغة العربية لغة ثانية؛ ففي هذا المجال عدد من الكفايات اللازمة التي يجب على المعلم أن يكون متمكنًا من الأداء الجيد من خلالها سواء كانت مهنية تدريسية أو معرفية أكاديمية أو ثقافية تواصلية.

إضافة لما سبق فهناك مهارات ينبغي أن يتقنها المتعلم لتحقيق متطلبات القرن الحادي والعشرين، مما يستدعي سلفًا أن يكون المعلم مؤهلًا تأهيلًا مناسبًا يسهم في بناء مهارات المتعلمين وبما يتناسب مع متطلبات هذا القرن، انطلاقًا من الإعداد الجيد للمعلم، وحسن اختياره. (شعيب، ٢٠٢٠: ٢)، وهذا مما يؤكد على ضرورة الاختيار وفقًا لأسس علمية؛ لتحقيق هذه المستهدفات، ولعل أحد هذه الأسس العلمية إجراء اختبار معياري لقياس مدى تمكن المعلم من الكفايات اللازمة لتعليم اللغة العربية لغة ثانية، وتحقيق الأهداف المنشودة لتنمية مهارات المتعلمين اللغوية والمهارية.

وأضاف (الحديبي، ٢٠٢٠: ١٣) أن تعليم اللغة العربية للناطقين بلغات أخرى يتطلب توفر عدد من الكفايات في المعلم الذي يقوم بتعليمها؛ ليكون عنصرًا فاعلًا في العملية التعليمية وبصورة تجعل المتعلمين:

- يقبلون على تعلم اللغة العربية.
- يحققون نواتج التعلم المطلوبة في المجالات المختلفة.
- يمتلكون مهارات التفكير اللازمة.
- يتعاملون مع المحتوى الرقمي والإلكتروني بكفاءة.
- يطبقون ما تعلموه في تواصلهم مع الناطقين باللغة العربية.

كفايات معلم اللغة العربية لغة ثانية:

الكفاية المهنية: ويقصد بها: مجموعة الخبرات الأدائية التي يمارسها معلم اللغة العربية لغة ثانية أثناء تعليمها، تتعلق هذه الأداءات بمراحل العملية التعليمية: التخطيط، والتنفيذ، والتقويم، من أجل تحقيق نواتج التعلم المستهدفة، باستخدام أحدث الأساليب في تعليم اللغات، ولبناء شخصية المتعلمين وتمكينهم من استخدام اللغة العربية في المواقف الاتصالية التي يطمحون إليها.

إضافة إلى أن تحسين كفايات معلم اللغة، سيكون له أثر في مساعدة المتعلمين الناطقين بلغات أخرى على تحقيق أهدافهم التعليمية، وهذا من نتائج الإعداد الجيد والتأهيل المهني لمعلمي اللغة. (Quarterlym، ٢٠٢٢)

فالكفاية المهنية لمعلمي اللغة العربية لغة ثانية تتطلب من المعلم في كافة مراحلها البحث عن أفضل الأساليب وأحدث التوجهات والممارسات التدريسية الفاعلة؛ لتحقيق نواتج التعلم المنشودة، ومراعاة قدرات المتعلمين ومستوياتهم. فقد ذكر صابر أن هذه الكفاية ضرورية لتحقيق أهداف تعليم اللغة، وتنمية مهارات المتعلمين اللغوية والمهارية. (صابر، ٢٠١٦: ٩٨٠).

إضافة لما سبق فإن الكفاية المهنية للمعلم تتطلب معرفته بطبيعة المهنة التي ينتمي إليها وأسسها الرئيسة، والقدرة على التخطيط والتنفيذ والتقويم

بما يناسب قدرات المتعلمين ومستوياتهم اللغوية. (البار؛ والقرني، ٢٠١٩: ١٥٠٨)، ويواكب التطورات الحديثة في مجال تعليم اللغات، فبناء الكفاية المهنية لدى معلمي اللغة العربية لغة ثانية أثناء الإعداد المهني لهم، وقياس مدى تمكنهم منها قبل ممارسة تعليمها جدير بالاهتمام، إضافة إلى اعتباره معيارًا للتعيين وممارسة تدريسها من خلال معلمين أكفاء.

الكفاية اللغوية: ويقصد بها: مجموعة الخبرات النظرية في عناصر اللغة العربية ومهاراتها وأساليبها التواصلية التي يجب أن يتقنها معلم اللغة العربية لغة ثانية ويوظفها؛ لتمكين المتعلمين من تعلم اللغة العربية بقواعدها الصحيحة نطقًا وكتابة وأسلوبًا وفهمًا ودلالة...، من أجل تحقيق نواتج التعلم المستهدفة في البرنامج التعليمي، والتركيز على الجوانب الوظيفية بما يتناسب مع نتائج دراسات علم اللغة التطبيقي التي تتعلق بتعليم اللغات، وفق أحدث الأساليب العلمية والتعليمية.

وتتمثل الكفاية اللغوية للمعلم في الطلاقة اللغوية في اللغة العربية نطقًا وقراءة وكتابة وتحدثًا، إضافة إلى الإلمام بقواعد اللغة نحوها وصرفها وأساليبها، والقدرة على تحليل وتذوق النصوص الأدبية. (صابر، ٢٠١٦: ٩٨٠).

فالخلاصة مما سبق أنّ بناء الكفاية اللغوية لدى معلمي اللغة العربية لغة ثانية وتمكنهم من تقديمها وفق قواعدها الصحيحة أحد أهم الكفايات التي ينبغي مراعاتها عند اختيارهم وترشيحهم لتعليمها؛ لبناء مهارات المتعلمين اللغوية والمهارية.

إضافة لما سبق فإن معرفة المعلم وإدراكه أهمية اللغة العربية، وقدرته على استخدامها بشكل صحيح، وتمكنه من مهاراتها وعناصرها، ومعرفته بأساليبها متطلبات أساسية لتعليمها لغة ثانية. (البار؛ القرني، ٢٠١٩: ١٥٠٧)

من هنا يرى الباحث أن معلم اللغة الثانية يجب أن يكون ملمًّا بمهاراتها وعناصرها وقواعدها النظرية والتطبيقية، إضافة إلى معرفته بما وصل إليه علم اللغة التطبيقي في مجال تعليم اللغات؛ لتحقيق مخرجات التعلم المنشودة وفق أحدث الأساليب وأنجعها.

الكفاية الثقافية: اللغة وعاء الثقافة، وليس من اليسير تعليم لغة ما دون التعرض لثقافة أصحابها (قيمهم واتجاهاتهم وأنماط معيشتهم وعقائدهم)، فالثقافة لها مكانة مهمة في تعليم اللغات وتعلمها، وتُعَد مكونًا أساسيًّا ومكملًا مهمًّا لتعليمها. فقد أثبتت الدراسات أن المعلومات والمعارف الثقافية هدف أساسي من أهداف أي مادة تعليمية لتعليم اللغات الأجنبية، وعامل مهم من عوامل النجاح في تعلم اللغة واستخدامها. (الفوزان، ٢٠١٧: ١٨٢)،

علاوة على ما سبق فإن التأهيل الثقافي لمعلم اللغة العربية لغة ثانية ضرورة ملحة؛ لذلك فإن قياس قدراته الثقافية اللازمة لتعليم اللغة العربية ومدى تمكنه من الأداء المناسب في مجال تعليمها، قد يسهم في بناء المهارات التواصلية بصورة جيدة ومناسبة للمواقف الحياتية للمتعلم.

ويقصد بالكفاية الثقافية: مجموعة من الخبرات التي يكتسبها معلم اللغة العربية عن طريق الممارسة المباشرة مع المجتمع اللغوي أو التأهيل الأكاديمي، من خلال البرامج التعليمية في الثقافة العربية والإسلامية والعالمية؛ من أجل التمكن من الأداء المناسب في المواقف التعليمية، والتعامل مع الثقافات الأخرى دون تحيز أو اصطدام.

فمعلم اللغة العربية لغة ثانية لابد أن يمتلك الكفايات الثقافية التي تؤهله لأداء عمله على الوجه اللائق، وأن هذه الكفايات يتعلق بعضها بالثقافة العربية والإسلامية، وبعضها الآخر يتعلق بالثقافة العالمية. (علي، ٢٠٢٢: ٧٣٣)

فمن الضروري تعزيز كفاءة التواصل بين الثقافات للمعلمين حول المفاهيم والظواهر الثقافية مع المتحدثين الأصليين للغة وغير الناطقين بها، والتي من

الضروري أن يمتلكها معلمو اللغات الأجنبية في المستقبل كي يستطيعوا تنمية مهارات المتعلمين، لا سيما في مرحلة مبكرة من عملية التعلم الخاصة بهم، وفهم ثقافة البلد الذي يدرسون فيه وثقافة اللغة التي يعلمونها. (Elizabeth، ٢٠٢١). ويؤكد الباحث على ضرورة مراعاة هذه الكفايات عند ترشيح المعلمين لتعليم اللغة العربية لغة ثانية.

وتعد الكفايات الثقافية لازمة للتعليم والتعلم، لاسيما في القرن الحادي والعشرين؛ حيث تؤكد عليها كثير من المعايير الدولية للمعلمين ووثائق الاعتماد، فيجب أن يكون المعلمون أكفاء ثقافيًّا حتى يتمكنوا من تلبية التنوع في فصولهم الدراسية وإعداد طلابهم إعدادًا مناسبًا للحياة ومتطلبات العصر. (Macqueen & Ferguson، 2020)

مما سبق يستنتج الباحث ضرورة أن يكون معلم اللغة العربية لغة ثانية قادرًا على أن يوظف الجوانب الثقافية والتواصلية في تعليمها لغة ثانية، إضافة إلى الممارسات التدريسية التي يطبقها أثناء تعليم اللغة، وللتأكد من مناسبة المرشح للعمل في مجال تعليم اللغة العربية لغة ثانية يرى الباحث ضرورة بناء اختبار معياري لقياس الكفايات اللازمة في معلم اللغة العربية لغة ثانية.

المحور الثاني: الاختبار المعياري لقياس كفايات معلمي اللغة العربية لغة ثانية:

الاختبار: أحد الوسائل التي يمكن من خلالها الحصول على بيانات تستخدم في القياس والتقويم أو التقييم. (الدهماني؛ وآخرون، ٢٠١٥: ١٨٨)

المعيار: (ع ي ر)، مِعْيار [مفرد]: ج: معايِيرُ: عيار؛ مقياسٌ يُقاسُ به غيرُه للحكم والتَّقييم «اخترته حسب معايير معيّنة - معيار الذَّهب / العيش - اختاروا الموظَّفين. غير مِعياريّ: مختلف أو غير ملتزم بمعيار معيَّن. (عمر، مادة: ع ي ر، ٣٥٠٦: ١٥٨٢).

المعيار في الاصطلاح: اتفقت عدة مصادر على تعريفه على ضوء نقطتين هما: ما يجب أن يعرفه المعلم، وما يكون قادرًا على أدائه. (الحديبي، ٢٠٠٨: ٢٠)

يعرفه الباحث إجرائيًا بأنه: جمل خبرية تصف ما ينبغي أن يعرفه المعلم ويوظفه من كفايات: مهنية، ولغوية، وثقافية، في تعليم اللغة العربية لغة ثانية؛ نتيجة لتأهيله المهني واللغوي والثقافي، ويتكون المعيار من عدد من مؤشرات الأداء التي يقوم بها المعلم أثناء تعليمه للغة العربية لغة ثانية.

ويقصد بالاختبار المعياري في هذا البحث: مجموعة من الأسئلة الاختبارية التي تستهدف قياس الكفايات الإجرائية لمعلمي اللغة العربية لغة ثانية؛ ليكون مرجعية علمية معتمدة في ترشيح معلمي اللغة العربية لغة ثانية، ويجب أن تقيس هذه الأسئلة الكفايات الإجرائية للمعلم ومدى قدرته على توظيفها أثناء تعليم اللغة العربية لغة ثانية.

خصائص الاختبار المعياري المقترحة:

- يحتوي على عدد من الأسئلة التي تقيس الكفايات الإجرائية التي توصل إليها البحث.
- يقيس الكفايات: المهنية، اللغوية، الثقافية لمعلم اللغة العربية لغة ثانية.
- يشمل الكفايات الإجرائية اللازمة لمعلمي اللغة العربية لغة ثانية.
- يمكِّن من الحكم على المختبَر من خلال درجته في الاختبار وعلى مدى مناسبته لتعليم اللغة العربية لغة ثانية.
- يصنف المختبَرين بناء على درجاتهم التي حصلوا عليها.
- يُبنى على أسس علمية تتمثل في تحكيمه، وتجريبه، وتحديد الكفايات الإجرائية التي يقيسها بعد التأكد من مناسبتها لمجال تعليم اللغة العربية لغة ثانية.
- يتصف الاختبار المعياري المقترح لقياس كفايات معلمي اللغة العربية لغة ثانية بالصدق والثبات والموضوعية وسهولة التطبيق.

ثانيًا: الدراسات السابقة:

دراسة: (الربابعة، ٢٠١٦)، بعنوان: «الكفايات التعليمية اللازمة لمعلمي اللغة العربية للناطقين بغيرها ومدى ممارستهم لها»، والتي هدفت إلى: بناء قائمة الكفايات التعليمية اللازمة لمعلمي اللغة العربية للناطقين بغيرها وبيان مدى ممارستهم لها، واستخدمت الدراسة المنهج الوصفي التحليلي لمناسبته لأهداف الدراسة، وتكونت عينة الدراسة من (٢٤) معلمًا ومعلمة، يدرسون اللغة العربية للناطقين بغيرها في مركز اللغات في الجامعة الأردنية، وتوصلت الدراسة إلى: بناء قائمة بالكفايات التعليمية اللازمة لمعلمي اللغة العربية للناطقين بغيرها. كما توصلت إلى وجود فروق ذات دلالة إحصائية في درجة ممارسة معلمي اللغة العربية للناطقين بغيرها للكفايات التعليمية، ويعزى ذلك للمؤهل العلمي. وأوصت الدراسة بتوجيه أنظار واضعي مناهج اللغة العربية للناطقين بغيرها إلى أهمية الكفايات التعليمية.

دراسة: (الحديبي، ٢٠١٦)، بعنوان: «فاعلية برنامج تدريبي قائم على معايير الجودة في تنمية معايير الأداء المهني لدى معلمي اللغة العربية للناطقين بلغات أخرى»، والتي هدفت إلى: تحديد معايير الأداء المهني اللازمة لمعلمي اللغة العربية للناطقين بلغات أخرى، وللوصول إلى نتائج الدراسة استخدم الباحث المنهج الوصفي والمنهج شبه التجريبي، وتكونت عينة الدراسة من (٢٠) معلمًا من معلمي اللغة العربية في معهد تعليم اللغة العربية لغير الناطقين بها في الجامعة الإسلامية بالمدينة المنورة، وتوصلت الدراسة إلى عدد من النتائج أهمها: قائمة بمعايير الأداء المهني اللازمة لمعلمي اللغة العربية للناطقين بلغات أخرى، تضمنت ثلاثة مجالات رئيسة انبثق عنها (٢٧) معيارًا، تفرع عنها (٢٠٧) مؤشرًا.

دراسة: (شعيب، ٢٠٢٠)، بعنوان: «معايير إعداد معلمي اللغة العربية للناطقين بلغات أخرى وفقا لمهارات القرن الحادي والعشرين»، التي هدفت

إلى إعداد قائمة بمعايير إعداد معلمي اللغة العربية للناطقين بلغات أخرى وفقا لمهارات القرن الحادي والعشرين، واستخدمت المنهج الوصفي التحليلي، وتوصلت إلى: معايير إعداد معلمي اللغة العربية للناطقين بلغات أخرى وفقا لمهارات القرن الحادي والعشرين؛ جاءت في ثلاثة مجالات هي: مجال الإبداع والابتكار، ومجال: مهارات المعلومات ووسائل الإعلام والتكنولوجيا، ومجال: مهارات الحياة والمهنة وتضم هذه المجالات عددًا من المهارات الفرعية انبثق عنها: (٥٣) مؤشرًا.

دراسة: (الحديبي، ٢٠٢٠)، بعنوان: «المعايير المعاصرة لإعداد معلم اللغة العربية للناطقين بلغات أخرى»، والتي هدفت إلى: إعداد قائمة بمعايير معلم اللغة العربية للناطقين بلغات أخرى في ضوء المعايير المعاصرة لإعداد معلم اللغات الأجنبية، واستخدم الباحث المنهج الوصفي للوصول إلى النتائج، وتوصلت الدراسة إلى قائمة بمعايير إعداد معلم اللغة العربية للناطقين بلغات أخرى على ضوء المعايير المعاصرة لإعداد معلم اللغات الأجنبية تضمنت: (٣) محاور رئيسة، انبثق عنها (١٢) مجالًا، تفرع عنها (٥٦) معيارًا، قسمت إلى (٦٠٢) من مؤشرات الأداء.

دراسة: (النوري، ٢٠٢٢)، بعنوان: «تصور مقترح لإعداد معلم اللغة العربية لغير الناطقين بها في ضوء التوجهات الحديثة ومتطلبات العصر الرقمي من وجهة نظر أعضاء هيئة تدريس اللغة العربية لغير الناطقين بها بمعهد تعليم اللغات بجامعة تبوك» والتي هدفت إلى: بناء تصور مقترح لإعداد معلم اللغة العربية لغير الناطقين بها في ضوء التوجهات الحديثة ومتطلبات العصر الرقمي من وجهة نظر أعضاء هيئة تدريس اللغة العربية لغير الناطقين بها بمعهد تعليم اللغات بجامعة تبوك، واستخدم الباحث المنهج الوصفي، وتوصلت الدراسة إلى أن درجة واقع إعداد معلم اللغة العربية لغير الناطقين بها في ضوء

التوجهات الحديثة كانت بمتوسط وقدره (٤,٢٧)، وأن درجة إعداد معلم اللغة العربية لغير الناطقين بها في ضوء متطلبات العصر الرقمي كانت بدرجة (٤,١٨).

دراسة: (Benjawan ؛Jira & Alongkorn 2022)، بعنوان: «نموذج لإعداد المدرسين لامتحان المعلم والترخيص حسب معايير الكفاءة لمجلس المعلمين في تايلاند»، والتي هدفت إلى: تقديم نموذج لإعداد المعلمين للامتحان وتطويره؛ للحصول على ترخيص مدرس وفقًا لمعايير الكفاءة الخاصة بمجلس المعلمين في تايلاند، استخدم الباحثون مزيجًا من الأساليب الكمية والنوعية، وتوصلت الدراسة إلى: حاجة المعلمين إلى الاستعداد للاختبار وفقا لمعايير كفاءة مجلس المعلمين في تايلند، وعلى مستوى عالٍ في جميع الموضوعات ذات الصلة.

التعقيب على الدراسات السابقة:

بعد اطلاع الباحث على هذه الدراسات اتضح أنَّها جميعا تؤكد على ضرورة الإعداد الجيد لمعلم اللغة العربية لغة ثانية بما يتناسب مع التوجهات الحديثة في تعليم اللغات، وبما يتفق مع أحدث الأساليب والإستراتيجيات ومتطلبات القرن الحادي والعشرين.

وعملت بعض هذه الدراسات على بناء قائمة بالمعايير اللازمة لمعلم اللغة من خلال التدريب على رأس العمل، مثل دراسة: (الحديبي، ٢٠١٦)، أو التأهيل الأكاديمي لمعلم اللغة العربية لغة ثانية؛ من هنا يرى الباحث ضرورة أن يكون الترشيح لتعليم اللغة العربية لغة ثانية وفقا لمعايير علمية تتضمن الكفايات: المهنية واللغوية والثقافية اللازمة، ولعل نتيجة المتقدم في الاختبار -إذا تم بناؤه بشكل جيد- تكشف عن مدى تمكنه من هذه الكفايات قبل التعيين.

وقد اتفق البحث الحالي مع بعض الدراسات السابقة في المنهج المستخدم للوصول للكفايات اللازمة. كما اتفق البحث الحالي مع دراسة: (الربابعة، ٢٠١٦)، ودراسة:

(الحديبي، ٢٠١٦)، ودراسة: (شعيب، ٢٠٢٠)، ودراسة: (الحديبي، ٢٠٢٠)، في عدد من الكفايات اللازمة لمعلم اللغة العربية لغة ثانية، ومهارات القرن الحادي والعشرين.

واختلف البحث الحالي عن الدراسات السابقة في إعداد تصور مقترح لبناء اختبار معياري لقياس الكفايات اللازمة لمعلمي اللغة العربية لغة ثانية، واستفاد البحث الحالي من الدراسات السابقة في إعداد قائمة الكفايات اللازمة لمعلمي اللغة العربية لغة ثانية، ويرى الباحث أن هذه القائمة تتضمن أهم الكفايات التي ينبغي توافرها في معلمي اللغة العربية لغة ثانية، وفق أحدث التوجهات والأساليب في تعليم اللغات.

إجراءات البحث:

منهجية البحث:

يستخدم البحث المنهج الوصفي، وذلك لتحديد الكفايات اللازمة لمعلمي اللغة العربية لغة ثانية، والخلفية النظرية للبحث. والمنهج الوصفي يطبق لمعرفة بعض الحقائق التفصيلية عن واقع الظاهرة محل الدراسة، مما يمكن الباحث من تقديم وصف شامل وتشخيص دقيق لذلك الواقع. (العساف، ٢٠١٢: ١٨٠)

إعداد أدوات البحث:

لتحقيق أهداف البحث أعدَّ الباحث قائمة بالكفايات اللازمة لمعلمي اللغة العربية لغة ثانية:

١. الهدف من إعداد القائمة:

يهدف الباحث من إعداد هذه القائمة إلى تحديد الكفايات المهنية واللغوية والثقافية اللازمة لمعلمي اللغة العربية لغة ثانية، وذلك للتركيز عليها في بناء اختبار كفايات معلمي اللغة العربية لغة ثانية.

٢. **مصادر إعداد القائمة:**

- مراجعة بعض الدراسات التي تناولت كفايات معلمي اللغة العربية لغة ثانية ومنها دراسة: (الربابعة، ٢٠١٦)، ودراسة:(الحديبي:٢٠٢٠)، ودراسة:(شعيب:٢٠٢٠).

- مراجعة بعض الكتابات النظرية، ونتائج بعض الدراسات السابقة ذات الصلة، ومنها: كفايات معلمي اللغة العربية: وزارة التربية والتعليم – الإدارة العامة للتقويم والجودة التربوية، وكتاب: المعايير والمسارات المهنية للمعلمين في المملكة العربية السعودية.

٣. **إعداد القائمة في صورتها الأولية:**

أُعدت القائمة في صورتها الأولية؛ لتحديد الكفايات اللازمة لمعلمي اللغة العربية لغة ثانية، وفقا للخطوات الآتية:

- كتابة الكفايات التي توصل إليها الباحث، من خلال مراجعة الدراسات السابقة والكتب المختصة في ذلك.

- عرض القائمة في صورتها الأولية على المحكمين، وقد اشتملت:
  - مقدمة توضح للمحكمين الهدف من إعداد القائمة.
  - المطلوب من المحكمين إبداء الرأي فيه.
  - كيفية تدوين الاستجابات التي تتناسب ورأي المحكم.
  - التعريف الإجرائي لكفايات معلمي اللغة العربية لغة ثانية.

اشتملت القائمة في صورتها الأولية على ثلاث كفايات رئيسة، انبثق عنها (١٢) كفاية فرعية، انبثق عنها (١٨٠) كفاية إجرائية، ويتضح ذلك من خلال الجدول الآتي:

جدول رقم (١) توصيف قائمة الكفايات اللازمة لمعلمي اللغة العربية لغة ثانية في صورتها الأولية:

م	الكفايات الرئيسة	الكفايات الفرعية		الكفايات الإجرائية
١	يتوفر لدى المعلم عدد من الكفايات المهنية التي تؤهله لتعليم اللغة العربية لغة ثانية.	٤		٦٤
٢	يتوفر لدى المعلم عدد من الكفايات اللغوية التي تؤهله لتعليم اللغة العربية لغة ثانية.	٤		٦٤
٣	يتوفر لدى المعلم عدد من الكفايات الثقافية التي تؤهله لتعليم اللغة العربية لغة ثانية.	٤		٥٢
	المجموع	٣	١٢	١٨٠

١. **عرض القائمة في صورتها الأولية على المحكمين:**

وعُرضت القائمة في صورتها الأولية على عدد من المحكمين، المختصين في مجال تعليم اللغة العربية لغة ثانية، وأصول التربية، والمناهج وطرق التدريس، واللغويات التطبيقية، والقياس والتقويم، وعلم النفس التربوي، وذلك للاسترشاد بآرائهم للوصول للقائمة النهائية للكفايات اللازمة لمعلمي اللغة العربية لغة ثانية. وقد طلب الباحث من المحكمين إجراء التعديلات المناسبة، من حيث:

- اتساق الكفايات الفرعية والإجرائية مع الكفايات الرئيسة التي تنتمي إليها.
- مناسبة الكفايات المضمنة وشمولها لكفايات معلمي اللغة العربية لغة ثانية.
- إضافة ما يرونه مناسبًا من كفايات فرعية أو إجرائية لم ترد في القائمة، وكذلك حذف الكفايات التي يرونها لا تتناسب مع أهداف البحث.

٢. **تعديل القائمة وفقا لآراء المحكمين:**

بعد إجراء التعديلات التي أشار إليها المحكمون، أعدَّ الباحث القائمة النهائية للكفايات: المهنية واللغوية والثقافية اللازمة لمعلمي اللغة العربية لغة ثانية، تمثلت في ثلاث كفايات رئيسة انبثق عنها (١٢) كفاية فرعية، انبثق عنها (١٨٠) كفاية إجرائية، والجدول الآتي يوضح ذلك:

جدول (٢) توصيف قائمة الكفايات اللازمة لمعلمي اللغة العربية لغة ثانية في صورتها النهائية:

م	الكفايات الرئيسة	الكفايات الفرعية		الكفايات الإجرائية
١	يتوفر لدى المعلم عدد من الكفايات المهنية التي تؤهله لتعليم اللغة العربية لغة ثانية.	٤		٦٤
٢	يتوفر لدى المعلم عدد من الكفايات اللغوية التي تؤهله لتعليم اللغة العربية لغة ثانية.	٤		٦٤
٣	يتوفر لدى المعلم عدد من الكفايات الثقافية التي تؤهله لتعليم اللغة العربية لغة ثانية.	٤		٥٢
المجموع		٣	١٢	١٨٠

نتائج البحث:

للإجابة عن السؤال الأول من أسئلة البحث الذي نصه: «ما الكفايات اللازمة التي ينبغي توفرها في معلمي اللغة العربية لغة ثانية؟

توصل الباحث إلى قائمة بالكفايات اللازمة لمعلمي اللغة العربية لغة ثانية تضمنت ثلاث كفايات رئيسة.

الكفايات الرئيسة التي توصل إليها البحث:

- يتوفر لدى المعلم عدد من الكفايات المهنية التي تؤهله لتعليم اللغة العربية لغة ثانية. وتضمنت (٤) كفايات فرعية.
- يتوفر لدى المعلم عدد من الكفايات اللغوية التي تؤهله لتعليم اللغة العربية لغة ثانية. وتضمنت (٤) كفايات فرعية.
- يتوفر لدى المعلم عدد من الكفايات الثقافية التي تؤهله لتعليم اللغة العربية لغة ثانية. وتضمنت (٤) كفايات فرعية.

تمثلت الكفايات المهنية اللازمة لمعلمي اللغة العربية لغة ثانية في أربع كفايات فرعية، انبثق عنها (٦٤) كفاية إجرائية، جاءت على النحو الآتي:

- التمكن من تطبيق الجوانب التربوية اللازمة في تعليم اللغة العربية لغة ثانية وتضمنت (٢٢) كفايات إجرائية.
- التمكن من تطبيق الجوانب النفسية اللازمة في تعليم اللغة العربية لغة ثانية. وتضمنت (١٤) كفايات إجرائية.
- التمكن من استخدام الوسائل التقنية والمعلوماتية المناسبة في تعليم اللغة العربية لغة ثانية. وتضمنت (١٤) كفايات إجرائية.
- التمكن من تطبيق الإجراءات الإدارية والقيادية في إدارة صفه. وتضمنت (١٤) كفايات إجرائية.

وتمثلت الكفايات اللغوية اللازمة لمعلمي اللغة العربية لغة ثانية في أربع كفايات فرعية، انبثق عنها (٦٤) كفاية إجرائية، جاءت على النحو الآتي:

- التمكن من القواعد النحوية والصرفية في تدريس اللغة لغة ثانية. وتضمنت (١٦) كفايات إجرائية.
- التمكن من القواعد الإملائية والكتابية اللازمة لتعليم اللغة العربية لغة ثانية. وتضمنت (١٢) كفايات إجرائية.
- التمكن من توظيف الأساليب اللغوية والبلاغية في تعليم اللغة العربية لغة ثانية. وتضمنت (١٢) كفايات إجرائية.
- القدرة على تطبيق ما توصل إليه علم اللغة التطبيقي في تعليم اللغات. وتضمنت (٢٤) كفايات إجرائية.
- وتمثلت الكفايات الثقافية اللازمة لمعلمي اللغة العربية لغة ثانية في أربع كفايات فرعية، انبثق عنها (٥٢) كفاية إجرائية، جاءت على النحو الآتي:
- التمكن من الجوانب الثقافية: العربية، والإسلامية، والعالمية، اللازمة في تدريس اللغة العربية لغة ثانية. وتضمنت (١٩) كفايات إجرائية.
- التمكن من تطبيق المهارات التواصلية والتعاونية اللازمة لتدريس اللغة العربية لغة ثانية. وتضمنت (١٤) كفايات إجرائية.
- التمكن من المهارات الإعلامية والتكنولوجية اللازمة لتدريس اللغة العربية لغة ثانية. وتضمنت (٩) كفايات إجرائية.
- التمكن من المهارات الإنتاجية والمحاسبية اللازمة لتدريس اللغة العربية لغة ثانية. وتضمنت (١٠) كفايات إجرائية.

للإجابة عن السؤال الثاني من أسئلة البحث الذي نصه: «ما التصور المقترح لبناء اختبار معياري لقياس كفايات معلمي اللغة العربية لغة ثانية؟»

في هذا التصور يقدم الباحث وصفًا مفصلًا لإجراءات بناء اختبار معياري لقياس كفايات معلمي اللغة العربية لغة ثانية، من خلال خطوات علمية؛ لقياس الكفايات الإجرائية التي توصل إليها البحث الحالي، وإجراءات تطبيق الاختبار على المتقدمين للعمل في مجال تعليم اللغة العربية لغة ثانية.

وقد تم بناء هذا التصور المقترح للاختبار المعياري لقياس كفايات معلمي اللغة العربية لغة ثانية، بناءً على الكفايات الإجرائية التي توصل إليها البحث الحالي، بعد استطلاع آراء المختصين في تعليم اللغة العربية لغة ثانية، والمختصين في علم اللغة التطبيقي، والمناهج وطرائق التدريس، وبناء الاختبارات، والقياس والتقويم، للتأكد من مناسبة الكفايات التي توصل إليها الباحث لقياس كفايات معلمي اللغة العربية لغة ثانية.

عنوان التصور المقترح:

(الخطوات الإجرائية لبناء اختبار معياري لقياس كفايات معلمي اللغة العربية لغة ثانية).

رؤية التصور: مرجع معتمد لقياس كفايات معلمي اللغة العربية لغة ثانية.

رسالة التصور: تقديم اختبارات معتمدة وفق أفضل المعايير المهنية والممارسات التدريسية لقياس كفايات معلمي اللغة العربية لغة ثانية.

الهدف العام من التصور المقترح:

يهدف هذا التصور إلى بناء اختبار معياري لقياس كفايات معلمي اللغة العربية لغة ثانية يكون هذا الاختبار مرجعًا معتمدًا لقياس كفايات معلمي اللغة العربية لغة ثانية.

أهداف التصور:

تتلخص أهداف التصور المقترح في:

- وصف مفصل لإجراءات بناء اختبار كفايات معلمي اللغة العربية لغة ثانية.
- صياغة الإطار الفلسفي والنظري للتصور المقترح.
- تحديد آليات وخطط تنفيذ التصور المقترح لتنفيذه.
- تحديد أدوات ووسائل التصور المقترح.

أسس بناء التصور المقترح:

- يقوم هذا التصور على عدد من الكفايات الرئيسة والفرعية والإجرائية التي توصل إليها الباحث لقياس كفايات معلمي اللغة العربية لغة ثانية، على ضوء نتائج الدراسة التحليلية لاستطلاع الرأي ومراجعة الأدبيات السابقة.
- يراعي هذا التصور التكامل بين المعرفة والتطبيق، أو الجانب المعرفي والجانب المهاري، مع مراعاة الأسس العلمية في بناء الاختبارات، وتحديد الكفايات المهنية واللغوية والثقافية المستهدفة.
- يراعي التصور الكفايات الإجرائية التي يمكن أن تكون مرجعية معتمدة في اختيار معلمي اللغة العربية لغة ثانية، فمن خلالها يمكن أن يتم الاختيار للعمل والترشيح له في مجال تعليم اللغة العربية لغة ثانية.
- يراعي هذا التصور بناء الاختبار على أسس علمية منها: الصدق والثبات والموضوعية، وسهولة التطبيق، عند إجرائه على الفئة المستهدفة.

محتوى التصور المقترح:

تضمن محتوى التصور المقترح: مقدمة عامة للتعريف بالتصور، العنوان، الهدف العام من التصور، أسس بناء التصور، أهداف التصور، محتوى التصور، إجراءات تطبيق الاختبار المعياري لقياس كفايات معلمي اللغة العربية لغة ثانية.

يتضمن التصور المقترح الإجابة عن الأسئلة الآتية:

- لماذا هذا التصور؟
- ماذا يتضمن هذا التصور؟
- كيف يتم بناء الاختبار وفق هذا التصور؟
- كيف يتم التحقق من تحقيق أهدافه؟

أولا: لماذا هذا التصور؟

أهدافه: تم تحديد أهدافه لتشمل:

- **الأهداف العامة:** يسعى هذا التصور إلى بناء اختبار معياري لقياس الكفايات اللازمة لمعلمي اللغة العربية لغة ثانية، يتم بناء على نتائجه ترشيح معلمي اللغة العربية لغة ثانية، إضافة إلى اعتماد الكفايات التي توصل إليها البحث مرجعية معتمدة في بناء المقررات التعليمية في برامج تأهيل معلمي اللغة العربية لغة ثانية وتدريبهم.
- **الأهداف الخاصة:** يهدف هذا التصور إلى: تقديم إجراءات بناء اختبار معياري لقياس الكفايات اللازمة لمعلمي اللغة العربية لغة ثانية. وتقديم خطوات إعداد هذا الاختبار وتنفيذه وتطبيقه بناء على الكفايات الإجرائية والفرعية والرئيسة التي توصل إليها الباحث.

ثانيا: ماذا يتضمن التصور المقترح لتحقيق الأهداف التي عُرضت في السؤال السابق؟

يتضمـن التصـور المقـترح عـددًا مـن المعـارف والخـبرات والمهـارات الـتي يتوقـع مـن معلـم اللغـة العربيـة لغـة ثانيـة الإلمـام بهـا مـن خـلال اسـتجابته علـى هـذا الاختبـار ومـن خلال النتيجـة الـتي يحصـل عليهـا في الاختبـار، وقـد تـم اختيـار هـذا المقـترح وفقـا لمـا يـأتي:

- ملاءمة الاختبار للكفايات الإجرائية المستهدفة.
- مراعاة الجوانب التربوية واللغوية والثقافية والفنية عند بناء الاختبار.
- شمولية الاختبار لجميع المعارف والمهارات المراد قياسها.
- مراعـاة التنـوع والمرونـة في الأسـئلة المقدمـة بمـا يتـلاءم مـع الكفايـات المطلوبة في معلـمي اللغـة العربيـة لغـة ثانيـة.
- التنوع في أساليب التقويم.
- إعـداد محتـوى الاختبـار في ضـوء الكفايـات الـتي توصـل إليهـا البحـث وفقـا لمعايـير اختيـار علميـة دقيقـة.

ثالثا: كيف يطبق الاختبار وفقا لهذا التصور؟

يتـم تطبيـق الاختبـار المعيـاري لقيـاس كفايـات معلـمي اللغـة العربيـة لغـة ثانيـة مـن خـلال مراحـل يتـم في كل مرحلـة إجـراءات جـاءت علـى النحـو الآتي:

مراحل خطوات الاختبار المعياري وتطبيقه وفق هذا التصور:

- **التحليل:** في هـذه المرحلـة يتـم تحليـل الكفايـات اللازمـة لمعلـمي اللغـة العربيـة لغـة ثانيـة وقـد توصـل الباحـث إلى (١٨٠) كفايـة إجرائيـة في هـذا البحـث.
- **التصميـم:** في هـذه المرحلـة يتـم بنـاء الاختبـار في صورتـه الأوليـة بحيـث يتضمـن كل سـؤال قيـاس كفايـة إجرائيـة ممـا تـم التوصـل إليـه.

- **التحكيم:** في هذه المرحلة يتم عرض الاختبار على عدد من الخبراء في مجال القياس والتقويم والخبراء في مجال تعليم اللغة العربية لغة ثانية.
- **التجريب:** في هذه المرحلة يتم تجريب الاختبار على عينة استطلاعية للتأكد من صدقه وثباته.
- **التطبيق:** في هذه المرحلة يتم تطبيق الاختبار بشكل فعلي على المتقدمين لتعليم اللغة العربية لغة ثانية.
- **التصحيح:** ويمكن أن يكون آليًا أو يكون من خلال لجنة مكلفة بذلك.
- **التقويم:** يتابع القائمون على تطبيق الاختبار في هذه المرحلة مدى تحقق الأهداف المنشودة من خلال متابعة استجابات المتقدمين للتطوير والتحسين.

لتطبيق الاختبار المعياري لقياس الكفايات اللازمة لمعلمي اللغة العربية لغة ثانية يمكن اتباع الإجراءات التطبيقية السابقة التي تحقق -بعون الله- ضمان اختيار المعلم لتعليم اللغة العربية لغة ثانية على أساس علمي دقيق.

نوع الأسئلة الاختبارية المقترحة:

- الاختيار من متعدد.
- الصواب والخطأ.
- المزاوجة.
- التكملة.

ويمكن تطبيق أي نوع من أنواع الأسئلة التي تناسب آلية إجراء الاختبار حيث يمكن تطبيق الاختبار:

- **حضوريًّا:** من خلال لجان لإجراء الاختبار، ويستجيب المختبرون على ورقة تتضمن الأسئلة الاختبارية.
- **حضوريًّا محوسبًا:** من خلال إجرائه على أجهزة الحاسوب في قاعات وبإشراف لجان مراقبة.
- **عن بعد في مراكز محددة خارج المملكة:** من خلال إجرائه عبر الشبكة العالمية (الإنترنت)، في مراكز محددة لدى الجهات التي تحتاج إلى معلمين بإشراف من الجهة التي تقوم على الاختبار.
- **عن بعد:** من خلال موقع إلكتروني يدخل إليه المتقدم وفق آليات فنية وتقنية مناسبة.

يقترح الباحث تصنيف المختبرين بناء على نتائجهم في الاختبار المعياري لقياس كفايات معلمي اللغة العربية لغة ثانية، فيما يأتي:

- من (٩٠ – ١٠٠)، متميز.
- من (٨٠ – ٩٠)، متقدم.
- من (٧٠ – ٨٠)، متمكن.
- من (٦٠ – ٧٠)، مقبول.
- من (٥٠ – ٦٠)، يحتاج إلى تدريب مكثف قبل التعيين.
- أقل من (٥٠)، يحتاج إلى تأهيل لتعليم اللغة العربية لغة ثانية.

رابعا: كيف نتأكد من تحقيق التصور المقترح لأهدافه التي بني من أجلها؟

يمكن تطبيق هذا التصور من خلال بناء اختبار تجريبي يطبق على عينة من المتقدمين للتأكد من قياسه للكفايات المستهدفة في مجال تعليم اللغة العربية لغة ثانية، وتحكيمه من عدد من الخبراء في مجال القياس والتقويم ومجال تعليم اللغة العربية لغة ثانية.

مصادر إعداد التصور:

- النتائج التي توصل إليها البحث الحالي الخاصة بكفايات معلمي اللغة العربية لغة ثانية.
- نتائج الدراسات السابقة التي تناولت كفايات معلمي اللغة العربية لغة ثانية، ومعايير اختيار معلمي اللغة العربية لغة ثانية.
- الأدبيات التي تناولت الاختبارات المهنية والتخصصية في مجال تعليم اللغة العربية لغة ثانية.
- خبرة الباحث أثناء عمله في مجال تعليم اللغة العربية لغة ثانية، وملاحظاته أثناء تنفيذ البرامج التدريبية لمعلمي اللغة العربية لغة ثانية في بعض الدول؛ حيث لاحظ أنَّ بعض من يقوم بتدريس اللغة العربية لغة ثانية بحاجة إلى تطوير كفاياته المهنية، واللغوية، والثقافية.

توصيات البحث، ومقترحاته:

بناء على النتائج التي توصل إليها البحث الحالي، يقدم الباحث بعض التوصيات، ويقترح عددا من البحوث المكملة لهذا البحث، وذلك كما يأتي:

هذه الطبعة إهـداء من المجمع
ولا يُسمح بنشرها ورقياً أو تداولها تجارياً

التوصيات:

١. تدريب المعلمين وأعضاء هيئة التدريس على تطبيقات بناء الاختبارات المهنية وإجراءاتها.

٢. تضمين الكفايات الإجرائية التي توصل إليها البحث في البرامج التدريبية المقدمة لمعلمي اللغة العربية لغة ثانية.

٣. الإفادة من نتائج المختبرين في تحديد الاحتياجات التدريبية لمعلمي اللغة العربية لغة ثانية.

٤. تأليف دليل إجرائي مبني على الكفايات التي توصل إليها البحث للراغبين في تعليم اللغة العربية لغة ثانية.

٥. تطبيق التصور المقترح لبناء اختبار معياري لقياس كفايات معلمي اللغة العربية لغة ثانية.

المقترحات:

١. بناء دليل إجرائي لتعليم اللغة العربية لغة ثانية مبني على الكفايات الإجرائية التي توصل إليها البحث.

٢. فاعلية برنامج تعليمي قائم على الكفايات الإجرائية التي توصل إليها البحث.

٣. فاعلية برنامج تدريبي لمعلمي اللغة العربية لغة ثانية قائم على الكفايات الإجرائية التي توصل إليها البحث الحالي.

٤. فاعلية برنامج تدريبي لمعلمي اللغة العربية لغة ثانية قائم على الكفايات المهنية التي توصل إليها البحث.

المراجع:

- البار، ابتهال محمد؛ والقرني، نورة سعد، (٢٠١٩)، تعليمية النحو العربي على ضوء الكفايات اللازمة لمعلمي اللغة العربية، مجلة الدراسات العربية، كلية دار العلوم، جامعة إلمنيا، مصر، ص ص ١٤٩٧ - ١٥١٧

- حامد، غادة صبري؛ حسان، حسن محمد؛ صالح، عادل منصور، (٢٠١٤)، شروط ومتطلبات الترخيص للعمل بمهنة التعليم (دراسة ميدانية بمحافظة الدقهلية)، مجلة بحوث التربية النوعية، مصر.

- الحديبي، علي عبد المحسن، (٢٠١٧)، تصور مقترح للتغلب على المشكلات التي تواجه الجهات المعنية بتعليم اللغة العربية للناطقين بلغات أخرى، بحث منشور، مجلة كلية التربية جامعة أسيوط، مج ٣٣، ع١، ص ص ١ - ٥٦

- الحديبي، علي عبد المحسن، (٢٠٢٠)، المعايير المعاصرة لإعداد معلم اللغة العربية للناطقين بلغات أخرى، بحث منشور على الشبكة.

- الحديبي، علي عبد المحسن؛ قاسم، محمد جابر؛ الحجور، صالح عياد؛ شيخ، أحمد محمد، (٢٠١٧)، معايير تعليم اللغة العربية للناطقين بلغات أخرى، مركز الملك عبد الله الدولي لخدمة اللغة العربية، الرياض.

- الدهماني، دخيل الله بن محمد؛ الزهراني، محمد سعيد؛ الخليفة، حسن جعفر؛ مسلم، حسن أحمد؛ الحديبي، علي؛ العبيدي، خالد، تحرير: الزهراني، مرضي، (٢٠١٥)، معايير الأداء المهني لمعلمي اللغة العربية، ط١، الرياض.

- الراجحي، عبده، (١٤١١هـ)، علم اللغة التطبيقي وتعليم العربية، جامعة الإمام، الرياض.

- الربابعة، إبراهيم حسن، (٢٠١٦)، الكفايات التعليمية اللازمة لمعلمي اللغة العربية للناطقين بغيرها ومدى ممارستهم لها، عمادة البحث العلمي / الجامعة الأردنية، العلوم الإنسانية والاجتماعية، مجلد ٤٣، ملحق ٤، ص ص ١٦٥١ - ١٦٦٩

- شحاتة، حسن؛ والسمان، مروان، (٢٠١٢)، المرجع في تعليم اللغة العربية وتعلمها، ط١، مكتبة الدار العربية للكتاب، القاهرة.

- شحاته، حسن، (١٩٩٦)، تعليم اللغة العربية بين النظرية والتطبيق، ط٣، الدار المصرية اللبنانية، القاهرة.

- شعيب، أبو بكر عبد الله، (٢٠٢٠)، معايير إعداد معلمي اللغة العربية للناطقين بلغات أخرى وفقا لمهارات القرن الحادي والعشرين، بحث منشور في مؤتمر اللغة العربية الدولي الرابع بالشارقة الذي نظمه المركز التربوي للغة العربية لدول الخليج بالتعاون مع وزارة التربية والتعليم والجامعة القاسمية، المنعقد في: ٢٩- ٢٠٢٠/١/٣٠م.

- صابر، هاني، (٢٠١٦)، إعداد معلم اللغة العربية بين أهداف الطالب والواقع العملي، تجربة الباحث في برامج تأهيل المعلم، أبحاث معهد ابن سيناء للعلوم الإنسانية، منشورات مركز الملك عبد الله الدولي للغة العربية، ص ص ٩٧٥- ٩٨٩

- العساف، صالح محمد، (٢٠١٢م)، المدخل إلى البحث في العلوم السلوكية، ط٢، دار الزهراء للنشر والتوزيع، الرياض.

- عطا، إبراهيم محمد، (٢٠٠٦)، المرجع في تدريس اللغة العربية، ط٢، مركز الكتاب للنشر، القاهرة.

- علي، أحمد حسان، (٢٠٢٢)، مكانة الثقافة في تدريس اللغة العربية للناطقين بلغات أخرى، Arastırmaları Dergisi، (29)، 727-737. DOI: 10.29000/rumelide.1164884.

- أبو عمشة، خالد حسين؛ العلوي، محمد إسماعيلي؛ العمري، فاطمة؛ البطل، محمود؛ صنوبر، أحمد؛ الحربي، خالد؛ عبد الرحيم، رائد؛ الحدقي، إسلام يسري، (٢٠١٧) الدليل التدريبي في تدريس مهارات اللغة العربية وعناصرها للناطقين بغيرها النظرية والتطبيق، منشورات مركز الملك عبدالله الدولي لخدمة اللغة العربية، الرياض، المملكة العربية السعودية.

- الفوزان، عبد الرحمن إبراهيم، (٢٠١٥)، إضاءات لمعلمي اللغة العربية لغير الناطقين بها، ط٢، العربية للجميع، الرياض.

- مدكور، علي أحمد؛ طعيمة، رشدي أحمد؛ هريدي، إيمان أحمد، (٢٠١٠) المرجع في تعليم اللغة العربية للناطقين بلغات أخرى، ط١، القاهرة، دار الفكر العربي.

- النوري، أسماء حسن، (٢٠٢٢)، تصور مقترح لإعداد معلم اللغة العربية لغير الناطقين بها في ضوء التوجهات الحديثة ومتطلبات العصر الرقمي من وجهة نظر أعضاء هيئة تدريس اللغة العربية لغير الناطقين بها بمعهد تعليم اللغات بجامعة تبوك، International Conference in Arabic Festival (INCAFA) Departemen Sastra Arab Fakultas Sastra Universitas Negeri Malang Tahun

- هيئة تقويم التعليم والتدريب، (١٤٣٩)، المعايير والمسارات المهنية للمعلمين في المملكة العربية السعودية، مكتبة الملك فهد الوطنية، الرياض.

المراجع الإنجليزية:

- Benjawan Keesookpun, Jira Jitsupa & Alongkorn Koednat, (2022), Model of Preparing Teacher Students for the Examination for a Teacher License According to the Competency Criteria of the Teachers' Council of Thailand, Faculty of Education, Suan Dusit University, 295 Dusit, Bangkok, Thailand, Received: January 10, 2022 Accepted: February 20, 2022 Online Published: doi:10.5539/ies.v15n3p95 URL: https://doi.org/10.5539/ies.v15n3p95, www.eric.Com
- Perez-Gomez, Francisco; Cortes Coss, Dina Elizabeth, 2021, On Intercultural Communicative Competence: Student-Teachers' Accounts of Colombian Cultural Identity, English Language Teaching, v14 n12 p69-90. www.eric.Com
- Macqueen, Suzanne Elizabeth; Reynolds, Ruth; Ferguson-Patrick, Kate, 2020, Investigating the Cultural Competence of Preservice Teachers: Comparisons and Considerations, Journal of International Social Studies, v10 n1 p113-137. www.eric.Com
- TESOL Quarterlym, 2022, A Journal for Teachers of English to Speakers of Other Languages and of Standard English as a Second Dialect, v56 n3 p934-960. www.eric.Com
- Gong, Yang Frank; Lai, Chun; Gao, Xuesong Andy, (2022), Language Teachers' Identity in Teaching Intercultural Communicative Competence. www.eric.Com

مواقع الشبكة العنكبوتية:

- عمر، أحمد مختار، معجم العربية المعاصرة، «ع ي ر»، ٣٥٠٦، تحت الرابط: https://shamela.ws/book/29511/19848 ، تاريخ الاسترجاع: ١٠/ ٧/ ١٤٤٤هـ، الساعة: ١٠,٤٥ مساء

- Wästlund, E., Reinikka, H., Norlander, T., & Archer, T. (2005). Effects of VDT and paper presentation on consumption and production of information: Psychological and physiological factors. Computers in Human Behavior, 21, 377–394
- Weir, C.J. (2005). Language Testing and Validation. New York, N.Y.: Palgrave Macmillan.
- Xie, Q. (2015). Do component weighting and testing method affect time management and approaches to test preparation? A study on the washback mechanism. System, 50, 56-68.
- Yeom, S. & Jun, S. (2020). Young Korean EFL Learners' Reading and Test-Taking Strategies in a Paper and a Computer-Based Reading Comprehension Tests. Language Assessment Quarterly, 17(3), 282-299.
- Ziefle, M. (1998). Effects of display resolution on visual performance. Human Factors, 40, 554–568.

- Sackstein, S., Spark, L. & Jenkins, A. (2015). Are e-books effective tools for learning? Reading speed and comprehension: iPad vs. paper. South African Journal of Education,35(4): 1-14.
- Sawaki, Y. (2001). Comparability of conventional and computerized tests of reading in a second language. Language Learning & Technology, 5, 38–59.
- Singer, L. M., & Alexander, P. A. (2017). Reading across mediums: Effects of reading digital and print texts on comprehension and calibration. The Journal of Experimental Education, 85(1), 155–172.
- Solak, E. (2014). Computer versus Paper-Based Reading: A Case Study in English Language Teaching Context. Mevlana International Journal of Education 4(1):202-211.
- Spiro, R. J., Klautke, H., and Johnson, A. K., (2015). All bets are off: How certain kinds of reading to learn on the web are totally different from what we learned from research on traditional text comprehension and learning from text. In. Spiro, R., DeSchryver, M., Hagerman, M., Morsink, P., and Thompson, P. (Eds.), Reading at a crossroads? Disjunctures and continuities in current conceptions and practices (pp. 45–50). New York: Routledge.
- Trackman, L. M., Alexander, P. A., and Berkowitz, L. B. (2019). Effects of Processing Time on Comprehension and Calibration in Print and Digital Mediums. The Journal of Experimental Education, 87(1), 101-115.
- Tsargari, C. (1994). Method effect on testing reading comprehension: How far can we go? M.A. Thesis, University of Lancaster. 96 pp. (Eric Document Reproduction Service No. Ed424768).
- Waiprakhon, T., and Jaturapitakkul, N. (2018). Test-Taking Strategies used in the Reading Section of the Test of English for Thai Engineers and Technologists: A Computer-based ESP Test. PASAA Journal, 55, 147-177.

- Kerr, M. A., & Symons, S. E. (2006). Computerized presentation of text: Effects on children's reading of informational material. Reading and Writing,19(1), 1–19.
- Nevo, N. (1989). Test-taking strategies on a multiple-choice test of reading comprehension. Language Testing. 6 (2) 199-215.
- Noyes, J. M., & Garland, K. J. (2003). VDT versus paper-based text: Reply to Mayes, Sims, and Koonce. International Journal of Industrial Ergonomics, 31(6), 411–423.
- Noyes, J. M., & Garland, K. J. (2008). Computer-vs. paper-based tasks: Are they equivalent? Ergonomics, 51(9), 1352–1375.
- Noyes, J. M., Garland, K. J., & Robbins, L. (2004). Paper-based versus computer-based assessment: Is workload another test mode effect? British Journal of Educational Technology, 35(1), 111–113.
- O'Malley, J. & Chamot, A. (1990). Learning strategies in second language acquisition. Cambridge: Cambridge University Press.
- Oxford, R. (1990). Language learning strategies: What every teacher should know. Boston: Heinle and Heinle.
- Park, H.; & and Kim, D. (2016). English Language Learners' Strategies for Reading Computer-Based Texts at Home and in School. Calico Journal, 33(3), 380–409
- Reinking, D. (1988). Computer-mediated text and comprehension differences: The role of reading time, reader preference, and estimation of learning. Reading Research Quarterly, 23, 484–498.
- Ross B., Pechenkina E., Aeschliman C., & Chase A.-M. (2017). Print versus digital texts: understanding the experimental research and challenging the dichotomies. Research in Learning Technology, 25. https://doi.org/10.25304/rlt.v25.1976

- Hamer, A. B., and McGrath, J. L. (2011). On-screen versus On-paper Reading: Students' Strategy Usage and Preferences. NADE Digest, 5(3). 26-39.
- Haddock, G. Foad. C, Saul, V. Brown, W. and Thompson, R (2020). The medium can influence the message: Print-based versus digital reading influences how people process different types of written information. British Journal of Psychology, 111, 443–459.
- Higgins, J., Russell, M., & Hoffmann, T. (2005). Examining the effect of computer-based passage presentation of reading test performance. The Journal of Technology, Learning and Assessment,3(4), 1–36.
- Hosseini, M., & Toroujeni, S. M. H. (2017). Replacing paper-based testing with an alternative for the assessment of Iranian undergraduate students: Administration mode effect on testing performance. International Journal of Language and Linguistics,5(3), 78.
- Hirano, K. (2009). Research on test-taking strategies in L2 reading. Bull. Joetsu Uni. Educ., 28(5): 157-165.
- Huang, H. D. (2016). Test-Taking Strategies in L2 Assessment: The Test of English for International Communication Speaking Test. Perceptual and Motor Skills, 123 (1) 64–90.
- Hughes, A. (1989). Testing for Language Teachers. Cambridge University Press, Cambridge.
- Jeong, H. (2012). A comparison of the influence of electronic books and paper books on reading comprehension, eye fatigue, and perception. The Electronic Library, 30(3), 390-408. https://doi.org/ 10.1108/02640471211241663

- Coiro, J., & Dobler, E. (2007). Exploring the online reading comprehension strategies used by sixth-grade skilled readers to search fir and locate information on the Internet. Reading Research Quarterly, 42 (2), 214–257.
- Creed, A., Dennis, I., & Newstead, S. (1987). Proof-reading on VDUs. Behaviour & Information Technology, 6, 3–13.
- Dillon, A. (1992). “Reading from Paper Versus Screens: A Critical Review of the Empirical Literature.” Ergonomics 35 (10): 1297–1326.
- Dudeney, G., & Hockly, N. (2016). Literacies, technology and language teaching. In F. Farr & L. Murray (Eds.), The Routledge Handbook of Language Learning and Technology (pp. 115–121). New York, NY: Routledge
- Duran, E., and O. Alevli. (2014). “The Effect of Reading on Screen on Comprehension of 8th Grade Students.” Research in Reading & Writing Instruction, 2(1): 1–11.
- Fulcher, G., & Davidson, F. (2007). Language testing and assessment. London, England: Routledge.
- Gable, T., & Thompson, C. (2011). Electronic readers might be viable education tools. iPad News Daily. Retrieved from: http://www.iPadnewsdaily.com/electronic-readers-might-beviable-education-tools-study-shows-1348/.
- Gough, P. B. (1972). One second of reading. In J. F. Kawanagh & I. G. Mattingley (Eds.), Language by ear and by eye. Cambridge, MA: MIT Press.
- Gu, X., Wu, B. & Xu, X. (2015). Design, development, and learning in e-Textbooks: what we learned and where we are going. Journal of Computers in Education. 2, 25–41.
- Güneş, F. (2009). “Efficiency in Screen Readings (Ekran Okumada Verimlilik).” Kalkınmada Anahtar Verimlilik Gazetesi 248: 26–28.

- Chen, H. Y. (2009). Online reading comprehension strategies among general and special education elementary and middle school students [Unpublished doctoral dissertation]. Michigan State University.
- Cheng, L. and DeLuca, C. (2011). Voices from test-takers: Further evidence for language assessment validation and use. Educational Assessment, 16(2), 104–122.
- Chifari, A., Chiazzese, G., Lucaiano, S., Merlo, G., Ottaviano, S., Allegra, M., (2010). A reflection on some critical aspects of online reading comprehension. Informatica, 34, 491–495.
- Chun, M. (1997). Research on text comprehension in multimedia environments. Language Learning & Technology. 1(1), 60–81.
- Cohen, A. D. (1996). Verbal reports as a source of insights into second language learners' strategies. Applied Language Learning, 7(1), 5–24.
- Cohen, A. D., & Upton, T. A. (2007). 'I want to go back to the text': Response strategies on the reading subtest of the new TOEFL. Language Testing, 24(2), 209–250.
- Cohen, A.D. (1984). On taking language tests: What the students report. Language Testing, 1(1), 70-81.
- Cohen, A.D. (1994). English for specific purpose in Brazil: The use of summary tasks. In C. Hill and K. Pary (Ed.), From testing to assessment: English as an international language. London: Longman.
- Coiro, J. (2003). Reading comprehension on the Internet: expanding our understanding of reading comprehension to encompass new literacies. The Reading Teacher, 56 (5), 458–464.

- Anderson, N. J. (2002). The role of metacognition in second/foreign language teaching and learning. ERIC Clearinghouse on Languages and Linguistics.
- Ackerman, R., & Lauterman, T. (2012). Taking reading comprehension exams on screen or on paper? A metacognitive analysis of learning texts under time pressure. Computers in Human Behavior, 28(5), 1816–1828.
- Aweiss, S. (1993). Comprehension assessment measures in second language and foreign language reading [Conference presentation]. Annual Meeting of the Central States Conference on the Teaching of Foreign Languages.
- Bachman, L. F. (1990). Fundamental considerations in language testing. Oxford University Press.
- Bachman, L. F. (2000). Modern language testing at the turn of the century: assuring that what we count counts. Language Testing. 17(1), 1–42.
- Baym, N. K. (2015). Personal connections in the digital age. John Wiley & Sons.
- Bernhardt, E. (1986). Cognitive processes in L2: an examination of reading behaviors. In: Lantolf, J. & Labarca, A. (Eds.), Research in second language learning: Focus on the classroom (pp. 35–51). Ablex.
- Bernhardt, E. B. (1990). A model of L2 text reconstruction: the recall of literary text by learners of Germany. In: LaBerge, A. (Ed.), Issues in L2: Theory as Practice, Practice as Theory. Ablex.
- Brun-Mercero, M. (2019). Online Reading Strategies for the Classroom. English Teaching Forum, 57(4), 2-11.
- Bvuum, V., Klu, E. K., and Adika, G.S. (2017). Test-taking strategies for second language: Implications for the classroom. Gender & Behavior, 15(2), 8909–8915.

References

- Abanomey, A. (2002). The effects of texts' authenticity on reading comprehension test-taking strategies used by adult Saudi learners of English as a foreign language [Unpublished doctoral dissertation]. Arizona State University, Arizona.
- Abanomey, A. (2012). Do EFL Saudi learners perform differently with online reading? An exploratory study. Journal of King Saud University – Languages and Translation, 25(1), 1–11.
- Abbott, M. L. (2006). ESL reading strategies: Differences in Arabic and Mandarin speaker test performance. Language Learning, 56(4), 633–670.
- Al-Amri, S. (2007). Computer-based vs. paper-based testing: Does the test administration mode matter? [Proceedings paper]. Meeting of the British Association of Applied Linguistics, Edinburgh, United Kingdom.
- Alderson, J. (1984). Reading in a foreign language: A reading problem or a language problem? In J. Alderson & A. Urquhart (Eds.), Reading in a foreign language (pp. 1–24). Longman.
- Alisaari, J., Turunen, T., Kajamies, A., Korpela, M., & Hurme, T. R. (2018). Reading comprehension in digital and printed texts. L1-Educational Studies in Language and Literature, 18, 1-18.
- Allan, A. (1992). Development and validation of a scale to measure test-wiseness in EFL/ESL reading test-takers. Language Testing, 9(2), 101–122.
- Anderson, J. (2001, May 11–12). Metacognitive awareness in language testing and teaching: Using think-aloud protocols [Conference presentation]. 3rd Annual Meeting, Midwest Association of Language Teachers, "Crossing Boundaries," University of Michigan, Anne Arbor, Michigan.

appropriate bottom-up test-taking strategies in the case of the second question – refute that argument. Participants would not have used more strategies and mobilized proper strategy types with screen reading if they had been under stress or experienced difficulties reading. In line with this, Noyes et al.'s (2004) observation that more effort and higher workload are involved with CBT in contrast to PBT may partially explain the significant application of more test-taking strategies with CBR than with PBR.

In conclusion, and as stated by Yeom and Jun (2020), "only a limited picture can be acquired by simply comparing scores on the two modes" (p. 282). In other words, a comprehensive understanding of the comparability of CBR and PBR cannot be gained by simply measuring students' performance with the two reading modes. In-depth examination of the processes involved in comprehending and answering questions, the effect of question type, and the affective orientation of test takers should top the agenda for future research. Such mechanisms must be thoroughly examined to define reading comprehension within this emerging and increasingly dominant context. In addition, test creators should be aware that the test design may encourage or discourage the use of specific test-taking strategies and that those strategies may have an effect on test-takers' performance, which in turn could affect the construct that the test is meant to measure. Overall, one cannot maximize the potential advantages of digital reading without in-depth knowledge, which can only be gained through extensive and focused research. Further studies are needed regarding some of the significant, yet under-researched, variables involved, such as processing speed, proficiency level, and test-wiseness strategies. The current investigation is hopefully a step in the right direction.

& Garland, 2003, 2008; Sawaki, 2001; Yeom & Jun 2020). The study outcomes support the assumption that a distinctive method of engaging in screen reading is required due to the accompanying new textual boundaries, linearity, and navigation style (Anderson, 2002; Chen, 2009; Coiro and Dobler, 2007; Solak, 2014). The significant difference in the mobilization of test-taking strategies across the two reading mediums supports Singer and Alexander's (2017) account that "such digital texts may place unique demands on readers' skilful and strategic processing not typically associated with the processing of printed text" (p. 155). McEneaney (2003) suggests that ''hypertext materials may be more cognitively demanding or require a greater degree of higher-level relational processing (Wenger and Payne, 1996) than traditional print'' (p. 2).

The study findings may also explain the results obtained by previous research (Abanomey, 2012; Duran & Alevli, 2014; Ross, et al., 2017) indicating that students show better performance with CBR than with PBR. The fact that the CB group, in a statistically significant manner, used more test-taking strategies may explain why some previous studies indicated that CB perform better. Compared with PBR, CB reading involves not only more strategies but also the proper use of the appropriate types of strategies based on the question and what it measures. The use of bottom-up strategies for the second question proves this. Screen reading seems to encourage test takers to look at and interact with the text more frequently in comparison to paper reading. This finding is consistent with Noyes and Garland's (2008) observation in their meta-analytic study that as technology advances, reading on screens becomes more natural. Recent studies tend to show a greater equivalence between CBR and PBR tests.

In addition, some researchers (Baym, 2015; Dudeney & Hockly, 2016; Güneş, 2009; Ziefle, 1998; Wästlund et al., 2005) suggest that reading from a screen is stressful, more difficult, and causes greater levels of tiredness and weariness than reading on paper. However, the study findings that statistically significantly more test-taking strategies were used with computer-based reading – and more

PB group and 52 test takers in the CB group used a combination of two strategies. For Question 4 (Table 12), however, test takers showed a somewhat different, though not statistically significant, manner. The question asked test takers to use a single strategy to guess the meaning of a word on the basis of the context in which it appeared. In comparison to 99 test takers from the CB group, 114 test takers from the PB group used one strategy. The table indicates that, in contrast to PB test takers, more CB test takers used three or more strategies to answer Question 4. The question of whether CB test takers face a difficulty with the screen size that forces them to use more than one strategy at one time should be addressed in future studies. Overall, the four tables show that there were no statistically significant differences between the two groups in terms of how they used their test-taking strategies. Two important variables to explore in the future are whether certain test-taking strategies are associated with correct answers and with which of the two reading formats, screen or paper, such test-taking strategies are used more.

Discussion

The use of CBR in EFL classes implies that new challenges will greatly impact L2 learners' ability to comprehend what they read, because of new dynamics and literacies that are assumed to emerge with this new format of reading (Chifari et al., 2010; Coiro 2003). Previous research has produced mixed results regarding the relationship between reading medium (print versus digital text) and reading comprehension, with some studies favouring the print format. The major goal of this study was to examine the possible variation in test-taking strategies that may result from the use of two reading formats, namely computer-based and paper-based.

A statistically significant finding is that the CB group mobilized more test-taking strategies than the PB group did. Moreover, compared with the PB group, participants in the CB group used statistically significantly more bottom-up test-taking strategies. The former finding lends support to the argument that interacting with CBR texts may involve particular skills, abilities, and cognitive processes that are somewhat different from those required in PBR reading (Dillon,1992; Noyes

Table 11. Chi-Square Test for the Difference Between the PB and CB Groups in the Number of Strategies Used to Answer Question 3

Number of Strategies	PB		CB		Total	
	No.	%	No.	%	No.	%
1	88	50.6	88	50.6	176	50.6
2	50	28.7	46	26.4	96	27.6
3 or more	36	20.7	40	23.0	76	21.8
Total	174	100.0	174	100.0	348	100.0
Chi-square (χ2)	0.377					
P-Value	0.828 (N.S.)					

Table 12. Chi-Square Test for the Difference Between the PB and CB Groups in the Number of Strategies Used to Answer Question 4

Number of Strategies	PB		CB		Total	
	No.	%	No.	%	No.	%
1	114	65.5	99	56.9	213	61.2
2	35	20.1	39	22.4	74	21.3
3 or more	25	14.4	36	20.7	61	17.5
Total	174	100.0	174	100.0	348	100.0
Chi-square (χ2)	3.256					
P-Value	0.196 (N.S.)					

Tables 9–12 indicate that test takers across the two groups tended to mobilize one strategy with each of the four questions in an almost identical fashion. For example, an equal number of test takers in the two groups (88) used only one strategy to answer Question 3. Similarly, to answer Question 2, 51 test takers in the

Tables 3–8 document the number of strategies used by the two groups with the different questions, either collectively or individually. Next, Tables 9–12 break down the test takers according to the number of strategies they used with each of the four questions: one, two, or three or more strategies.

Table 9. Chi-Square Test for the Difference Between the PB and CB Groups in the Number of Strategies Used to Answer Question 1

Number of Strategies	PB		CB		Total	
	No.	%	No.	%	No.	%
1	84	48.3	76	43.7	160	46.0
2	64	36.8	58	33.3	122	35.1
3 or more	26	14.9	40	23.0	66	19.0
Total	174	100.0	174	100.0	348	100.0
Chi-square (χ2)	3.665					
p-Value	0.160 (N.S.)					

Table 10. Chi-Square Test for the Difference Between the PB and CB Groups in the Number of Strategies Used to Answer Question 2

Number of Strategies	PB		CB		Total	
	No.	%	No.	%	No.	%
1	90	51.7	88	50.6	178	51.1
2	51	29.3	52	29.9	103	29.6
3 or more	33	19.0	34	19.5	67	19.3
Total	174	100.0	174	100.0	348	100.0
Chi-square (χ2)	0.047					
p-Value	0.977 (N.S.)					

Table 7. Z-Test for the Differences in the Means of Test-Taking Strategies Between the PB and CB Groups with Question 3

Strategies Type	PB		CB		Z-Test	
	No	%	No	%	Z value	Sig.
Top-down	161	50.2	160	49.8	0.079	(N.S.)
Bottom-up	155	49.5	158	50.5	0.240	(N.S.)
Total	316	49.8	318	50.2	0.112	(N.S.)

Table 8. Z-Test for the Differences in the Means of Test-Taking Strategies Between the PB and CB Groups with Question 4

Strategies Type	PB		CB		Z-Test	
	No	%	No	%	Z value	Sig.
Top-down	114	49.1	118	50.9	0.371	(N.S.)
Bottom-up	166	46.8	189	53.2	1.726	(N.S.)
Total	280	47.7	307	52.3	1.576	(N.S.)

Tables 5, 7, and 8 show the similarities in participants' use of test-taking strategies with Questions 1, 3, and 4, respectively. There were no statistically significant differences across the two groups in their use of both top-down and bottom-up test-taking strategies with the three aforementioned questions. Table 6, however, shows a noteworthy result. The CB group, with a Z value of 2.03 at a 0.05 level, used significantly more bottom-up test-taking strategies with Question 2 in contrast to the PB group. This question was presumed to be answered using bottom-up strategies, as it involved referring to a specific part of the text to identify referents of a pronoun. This finding may indicate that CBR encourages test takers to refer to the passage more frequently than PBR does. Ackerman, & Lauterman (2012) argued that there is more self-monitoring with reading texts digitally especially if the reading time is limited. This may also explain why some studies show that test takers perform better with CBT than with PBT.

Regarding the third question, which measured students' ability to scan the text for specific information, the two groups of participants used almost equal numbers of bottom-up and top-down test-taking strategies. Finally, with the fourth question, which asked students to guess the meaning of a vocabulary item, Table 4 indicates that the CB group tended to use more bottom-up test-taking strategies in comparison to the PB group. The two groups' use of top-down test-taking strategies is almost identical. In total, the CB group mobilized more overall test-taking strategies, with a marginal difference of 27 strategies. Tables 5–8 indicate whether the differences in the number of test-taking strategies with each of the four questions are statistically significant or not.

Table 5. Z-Test for the Differences in the Means of Test-Taking Strategies Between the PB and CB Groups with Question 1

Strategies Type	PB		CB		Z-Test	
	No	%	No	%	Z value	Sig.
Top-down	120	46.5	138	53.5	1.585	(N.S.)
Bottom-up	185	49.3	190	50.7	0.365	(N.S.)
Total	305	48.2	328	51.8	1.293	(N.S.)

Table 6. Z-Test for the Differences in the Means of Test-Taking Strategies Between the PB and CB Groups with Question 2

Strategies Type	PB		CB		Z-Test	
	No	%	No	%	Z value	Sig.
Top-down	193	52.4	175	47.6	1.327	(N.S.)
Bottom-up	116	45.5	139	54.5	2.037	0.05
Total	309	49.6	314	50.4	0.283	(N.S.)

Question Number	Strategies Type	PB		CB		Total	
		No	%	No	%	No	%
Question 3	Top-down	161	50.9	160	50.3	321	50.6
	Bottom-up	155	49.1	158	49.7	313	49.4
	Total	316	100.0	318	100.0	634	100.0
Question 4	Top-down	114	40.7	118	38.4	232	39.5
	Bottom-up	166	59.3	189	61.6	355	60.5
	Total	280	100.0	307	100.0	587	100.0

Table 4 details the number of test-taking strategies applied with each of the four questions. It reveals that, overall, participants across the two groups tended to use more bottom-up test-taking strategies with the first question. This questions measured their ability to determine the main idea of the passage. Participants were expected to use more top-down test-taking strategies because of the demand of the question itself; however, they mobilized 258 top-down strategies in contrast to 375 bottom-up test-taking strategies. Top-down strategies are usually associated with questions that require higher thinking skills, namely inferring, self-monitoring, evaluation and making judgments.

Similarly, for the second question, the participants used more top-down than the bottom-up test-taking strategies. Question 2 asked the students to identify the referent of a pronoun, and in such case it is expected that more bottom-up test-taking strategies be utilized. A significant feature of bottom-up reading comprehension test-taking strategies is the focus on words and grammatical structure identification. However, the two groups of students deployed a total of 368 top-down strategies and only 255 bottom-up strategies. Table 4 also indicates that the PB group was the source of the observable tendency to mobilize more top-down strategies (193 in comparison to 175 by the CB group). On the other hand, The CB group used more bottom-up test-taking strategies than the PB group did. This observation has considerable implications, as will be shown in Table 6.

Table 3. A Z-test of the Differences in the Means of Test-Taking Strategies between the PB and CB Groups with all Questions

Strategies Type	PB		CB		Z-Test	
	No	%	No	%	Z value	Sig.
Top-down	588	49.9	591	50.1	0.124	(N.S.)
Bottom-up	995	47.8	1,086	52.2	2.821	0.01
Total	1,583	48.6	1,677	51.4	2.328	0.05

Table 3 indicates that both groups showed statistically significant differences in their overall use of test-taking strategies. At the 0.05 level, the CB group employed significantly more test taking strategies in comparison to the PB group. Furthermore, the table shows that, with a Z value of 2.821 at the 0.01 level, the CB group was statistically significantly different from the PB group in the use of more bottom-up test-taking strategies. However, both groups showed no statically significant difference in their mobilization of top-down test-taking strategies. An important question that arises is whether such differences in the overall use of test-taking strategies are present in the test-taking strategies used with each of the four questions. The answer to this inquiry is presented next in Table 4.

Table 4. Total Number of Strategies with Each of the Four Questions

Question Number	Strategies Type	PB		CB		Total	
		No	%	No	%	No	%
Question 1	Top-down	120	39.3	138	42.1	258	40.8
	Bottom-up	185	60.7	190	57.9	375	59.2
	Total	305	100.0	328	100.0	633	100.0
Question 2	Top-down	193	62.5	175	55.7	368	59.1
	Bottom-up	116	37.5	139	44.3	255	40.9
	Total	309	100.0	314	100.0	623	100.0

Results

In line with study's overall goal of determining whether a change in the medium of reading presentation has any effect on the test-taking strategies used in a reading test, I present the results related to three major questions: (1) Does the medium of reading (PB and CB) affect the number of test-taking strategies? (2) Does the medium of reading (PB and CB) affect the manner in which Saudi EFL learners use reading comprehension test-taking strategies? (3) Does the medium of reading (PB and CB) affect the type of reading comprehension test-taking strategies used (top-down and bottom-up)?

Table 2. Total Number of Strategies Used in Answering the Questions with Both Formats

Strategies Type	PB		CB		Total	
	No	%	No	%	No	%
Top-down	588	37.1	591	35.2	1,179	36.2
Bottom-up	995	62.9	1,086	64.8	2,081	63.8
Total	1,583	100.0	1,677	100.0	3,260	100.0

Table 2 indicates that the participants used a total of 3,260 strategies, with both CB and PB reading formats. The table also shows that there was similarity in the total number of overall strategies used by the CB and PB groups: The CB group used 1,677 strategies to answer the four questions, while the PB group used 1,583 strategies. This similarity in number is also applicable to the two categories of strategies used (top-down and bottom-up): The PB group used 588 top-down and 995 bottom-up strategies, and the CB group used 591 top-down and 1,086 bottom-up strategies. In addition, Table 2 indicates that, overall, participants used more bottom-up than top-down strategies to answer the four questions, regardless of the presentation medium. The CB group, specifically, tended to use this type of test taking strategies more than the PB group. Whether these numbers are statistically significant or not is shown in Table 3 below.

with the same text but in a linear traditional paper-based format and then asked to answer the test questions on the same page. This procedure took place over five semesters, which aided me in collecting data from a large group of subjects.

Upon the completion of the first stage, the researcher utilized a design that encompassed a four-question reading comprehension test as well as a questionnaire on the test-taking strategies. In such a design, the reading passages were shown on the first page. The first question regarding the passage appeared on the second page, followed by a question about the strategies participants used to answer that question. The same list of inquiries regarding the test-taking strategies accompanied the second, third, and fourth comprehension questions on the third, fourth, and fifth pages, respectively. Such a design, which Abanomey (2002), Nevo (1989), and Tsagari (1994) proved to be a valid, rapid, and effective tool in collecting the valuable data, allowed for immediate feedback from test takers (Chadwick et al., 1989; Chamot, 2001; Forcese & Richer, 1973). In addition, no coaching occurred within the present research on test-taking strategies. In other words, participants were not given any instructions on test-taking strategies, as it may have raised their awareness and influenced them to utilize strategies that they would not otherwise have employed.

Lastly, all questions asking participants about their strategies in answering the reading test were written in the Arabic language. The list of strategies from which they could choose was also written in Arabic. This decision was made because it might be difficult for participants to write about their strategies in their L2. Cohen (1994) addresses the importance of using the test takers' native language "in order to eliminate the problem of the nonnatives' proficiency in the target language interfering with the reporting process" (p. 125).

ability to identify referents of pronouns as a means of examining the grammatical ability of test takers participating in this study. These higher-level and lower-level processes were examined simultaneously and within the same test. Four multiple-choice questions were included: The first question assessed the test takers' ability to skim the text to deduce the main idea. The second question was intended to investigate participants' grammatical ability in EFL. While the third question asked participants to scan the text to search for pieces of information, the fourth question tested participants' vocabulary knowledge.

The researcher prepared a test-taking strategies questionnaire for use with both test presentation modes. In accordance with the purpose of the current investigation, it was necessary to utilize a design that would incorporate both a test and an elicitation tool that could examine and report the test-taking strategies used by the participants in the study. The author employed a questionnaire consisting of a checklist of test-taking strategies as guided by findings and recommendations of previous studies (Abanomey, 2002; Cohen, 1984, 1994; Nevo, 1989; Tsagari, 1994). The list included two types of test-taking strategies: Bottom-up and Top-down. Bottom-up strategies are the lower-level reading processes used to construct meaning from the most basic units of language, including letters, letter clusters, and words. According to Gough (1972), in bottom-up model, the reading process proceeds in serial fashion, from letter to sound, to words, to grammatical structure identification to meaning. Bottom-up reading strategies begin with letter-sound correspondences (the bottom) to achieve comprehension (the top). Top-down strategies, on the other hand, are holistic and proceed from the whole to the parts. Students create meaning by using background knowledge, making predictions, and searching texts to confirm or deny predictions made.

The researcher presented the four abovementioned questions and the questionnaire to both the CB group (on computer screen) and the PB group (on separate sheets of paper). The CB group was asked to read a passage that appeared on computer screens in a language lab, while the PB group was presented

هذه الطبعة إهـداء من المجمـع
ولا يُسمح بنشرها ورقياً أو تداولها تجارياً

Materials and Procedures

The purpose of this study was to determine which, if any, variations in test-taking strategy occurred in a reading test when the text was presented in the CB format as opposed to the traditional PB mode of reading. The research question guiding the study was as follows: Do EFL test takers use different test-taking strategies when the text is presented on a computer screen than when the text is presented on paper?

Writing the multiple-choice reading questions for the current study could not have been done without first deciding on a general reading comprehension test framework. Another important consideration was ensuring that various questions tap into different sub-reading skills.

Writing test items also involved deciding on a theory upon which the test constructor based the test. In the current study, it was adhered to Bernhardt's (1986, 1990) second-language reading model as the main source of test specifications. In addition to assigning significant importance to both higher-level and lower-level processes, Bernhardt's (1986, 1990) model considers L2 vocabulary knowledge and L2 grammatical skills to be critical in facilitating second-language reading comprehension. Most significantly, the model allows for the interaction between text-based elements, such as phonemic/graphemic features, syntactic feature recognition, and word recognition, and extra text-based elements, such as textual perceptions, prior knowledge, and metacognition. The model "focuses on the connected interactions between various textual features and influences external to the text" (Aweiss, 1993, p. 18). As a result—and based on the theory underlying Bernhardt's (1986, 1990) second-language reading model and following Hughes's (1989) and Cohen's (1994) recommendations—it was essential to tap into two kinds of reading skills: macro and micro skills. On the one hand, macro skills included scanning the passage to locate specific pieces of information and skimming the passage to infer the main idea conveyed by the writer. Micro skills, on the other hand, included the ability to identify the meaning of unknown vocabulary items based on the surrounding context and the

Method

Research Questions

The main objective of this study was to determine whether a change in the medium through which a reading text is presented has any effect on the number and type of reading comprehension test-taking strategies used by Saudi EFL learners in reading comprehension tests. This inquiry could be answered by addressing the following three questions: (1) Does the medium of reading (PB and CB) affect the number of test-taking strategies? (2) Does the medium of reading (PB and CB) affect the manner in which Saudi EFL learners use reading comprehension test-taking strategies? (3) Does the medium of reading (PB and CB) affect the type of reading comprehension test-taking strategies used (top-down and bottom-up)?

Participants

A total of 348 undergraduate male students at Riyadh College of Technology participated in this study. At the time of the study, the students were enrolled in a scientific reading course as one of the requirements to complete their bachelor's degree. I chose the participants based on their proficiency level as determined by their scores after finishing a one-semester intensive English course: a score between 70 and 80, which the majority of students achieved. It should be noted that students who fail to score 70 are required to retake the intensive English course. I divided students into two groups: PBR and CBR. Table 1 shows the distribution of subjects in the present study.

Table 1. The Distribution of Students in this Study

Medium of Reading	No. of Students	Percentage
PB	174	50%
CB	174	50%
Total	348	100%

authors, the increased cognitive demands of the digital format may explain this affective orientation. Duran and Alevli (2014) included the variable of gender when examining the effect of screen reading on the comprehension of Turkish eighth grade students. The researchers found that after reading on screen, participants' reading comprehension scores were significantly higher than after reading on printed materials. They also found a significant difference caused by gender: Female students' reading comprehension scores were superior to those of male students on both reading formats.

The divergent results highlight the importance of carefully examining the various factors involved in reading on both formats within testing situations, with the most crucial factor being EFL learners' actual processes when engaging in paper and screen reading (Haddock et al., 2020; Kerr & Symons, 2006; Noyes & Garland, 2003; Reinking, 1988; Sawaki, 2001). Yeom and Jun (2020) accurately state that "few studies investigated the extent to which students engage in similar processes of reading in the two test-presentation modes" (p. 284). In line with this, there have been calls for further studies that examine the comparability of paper- and computer-based tests from various perspectives, including those that focus on "the process rather than product of reading" (Sawaki, 2001, p. 51).

Given the paucity of studies comparing the strategies EFL students apply to answer comprehension questions on reading texts that are presented in paper- and computer-based forms (Fulcher & Davidson, 2007; Higgins, et al., 2005; Hosseini & Toroujeni, 2017; Yeom & Jun, 2020), the purpose of this study is to determine whether any differences exist in the reading comprehension test-taking strategies mobilized with the two formats. To my knowledge, the available research comparing CBR to PBR does not control the two variables of reading format and test items format. That is, it is unclear whether the reported discrepancy in performance can be attributed to the reading presentation format (paper or screen) or to the test question format (printed paper-based test questions or computer-based test questions).

processes when responding to reading comprehension questions presented to them in paper and computer test-presentation modes. They found not only no statistically significant difference between the strategies used on the reading tests delivered in both formats but also "very few statistically significant differences in strategy use between the two modes" (Yeom and Jun, 2020, p. 282). Another significant finding is the reported preference among most participants to take the test on paper despite observed similarities in test performance and strategy use between the two formats. Such a finding can be attributed to the reported difficulties with the computer mode and participants' lack of familiarity with reading on a computer screen.

Examining the mode of presentation effect on cognitive workload with a group of undergraduate students at the University of Bristol, Noyes et al. (2004) found that "although there were no significant differences in the comprehension scores or the overall workload scores for paper and computer" (p.112), more effort and higher workload was involved with CBT than with PBT, especially with lower-performing test takers. In a study that compared information retention with printed and electronic texts, Gable and Thompson (2011) found no significant difference in the time needed to read the information. In addition, although reading electronic texts improved short-term memory retention, the authors observed no significant difference between the two forms in terms of their effect on long-term retention. Another noteworthy finding was the positive effect that familiarity with computers had on the speed of reading.

Wästlund et al. (2005) investigated the effect of CBR on Swedish readers' perceptions of written texts. In their study, the authors asked participants to read text in either a printed format or PDF format (on a computer) and assessed comprehension thereafter. In addition to the finding that participants who read the printed text achieved greater levels of comprehension than those who read the digital text, Wästlund et al. (2005) reported that participants found the digital format to be particularly stressful when compared to the paper format. To the

assessments than on computer-based tests," whereas Abanomey (2012) found that "students who were assigned the internet-based reading test showed a better capacity to answer the questions correctly in comparison to those who were asked to take the same test in a traditional print-based format" (p. 9). In their assessment of related works, Yeom and Jun (2020) elaborate on the difficulty of making "a direct comparison between these studies because of the variation in age of the participants, target language, instruments, and contexts" (P. 283). Çınar et al. (2021) concur with this argument and state that "the lack of homogeneity of the test groups used in experimental settings of those studies" (p. 258) leads to discrepancies across literature on the effect of reading format on performance.

Some research on CBR comprehension, albeit limited, assesses language learners' reading comprehension strategies and compares conventional PBR to CBR. Researchers propose a set of differences that distinguishes the two reading formats from each other (Anderson, 2002; Coiro and Dobler, 2007; Solak, 2014). Some of these differences "relate to textual boundaries, linearity, and navigation" (Chen, 2009, p. 34). While paper-based texts are usually linear, computer-based texts are often nonlinear or multilinear (Chen, 2009), which affects the reading path. In contrast to reading traditional PB texts, which often involves a predetermined and predictable path, reading a CB text involves a random and unpredictable route. Coiro (2003) argues that reading from computer screens makes comprehension more challenging in comparison to conventional PBR, in part due to the limited space on the computer screen through which the reader interacts with the text. Ziefle (1998) and Wästlund et al. (2005) indicate that computer screen reading causes greater levels of tiredness and is more stressful than PBR. Given that such disparities may affect reading comprehension in general and reading comprehension assessment in particular, studies examining the comparability of the two modes have emerged within the last 20 years.

Yeom and Jun (2020) examined whether young Korean EFL learners use similar

Available research on the use of CBR in comparison to PBR in reading comprehension assessment focuses mainly on whether the presentation format of reading texts influences students' performance, mostly in L1 reading. Investigations into the effect of paper- versus computer-based presentation modes on learners' performance have produced a range of somewhat inconsistent findings (Yeom & Jun, 2020; Solak, 2014; Genc 2011). Variation in the comprehension assessment measures and the use of different study designs could be part of the reason for such disparity in research findings (Alisaari et al., 2018). On the one hand, several studies report that PBR, in contrast to CBR, leads to better learners' performance, represented by test score (Alamri, 2007; Ebrahimi, 2020; Choi et al., 2003; Creed et al., 1987; Haddock et al., 2020; Hosseini & Toroujeni, 2017; Kerr & Symons, 2006; Kim & Huynh, 2008; Lauren et al., 2019; Mangen et al., 2013; Muter et al., 1982). Jeong (2012) found that the reading comprehension scores of the 56 participants were significantly better when reading printed texts compared to digital texts. In addition, students reported preferring printed texts over digital texts.

On the other hand, some studies report that digital reading leads to enhanced comprehension and better understanding than PBR (Abanomey, 2012; Aeschliman, & Chase, 2017; Duran and Alevli, 2014; Ross, Pechenkina; 0Singer & Alexander, 2017). Kerr and Symon (2006) found, for instance, that when the participants in their study (60 fifth grade students) read the digital texts, they had better comprehension scores in contrast to when they read the printed texts. Much research within this area, however, reports no significant difference between the two presentation formats (Alisaari et al., 2018; Higgins et al., 2005; Pommerich, 2004; Reinking, 1988; Srivastava & Gray, 2012).

However, few studies have compared the actual reading comprehension test-taking strategies used with both reading formats (Yeom & Jun, 2020). The observation of inconsistent findings is even reported within the same local context. According to Alamri (2007), Saudi students "tend to score better on paper-based

- opportunities" (p. 381). New methods of interaction with widely used electronic documents and
- hypertexts on computer screens add to the difficulties that computer-based readers face (Baym,
- 2015; Dudeney & Hockly, 2016; Güneş, 2009). For example, "moving between texts, pages, and
- sites puts a strain on cognitive resources" (Brun-Mercer, 2019, p. 2). In addition, eye-tracking
- research has indicated that other types of texts on the same page being read causes distractions
- and eye fatigue (Nielsen and Pernice 2010). Alisaari et al. (2018) state that "reading digital texts
- is speculated to be cognitively more challenging than reading printed texts because a backlit
- computer screen causes eye fatigue and possibly interferes with memory retrieval" p.5. To make
- a valid interpretation of computer based test scores, Yeom & Jun, 2020 call for conducting more
- research on the comparability between computer-based and paper-based reading to examine
- whether "a difference in the presentation mode could result in changes in the constructs that are measured through the test" (p. 282).

learners, and test users alike" (Sawaki, 2001, p. 38). However, "despite this increased use of information and communication technologies in school and at home, little is known about the reading patterns and cognitive processes readers use in nonlinear digital texts" (Coiro and Dobler, 2007, p. 217). The major objective of this paper is to examine whether variation exists in the test-taking strategies that Saudi EFL learners use with computer-based reading (CBR) and paper-based reading (PBR). Few studies have examined whether the presentation format of the passages used in reading comprehension tests (screen and paper) have any effect on the number and type of test-taking strategies and the manner in which such strategies are used.

Literature Review

- The increasingly frequent application and integration of computer-based technologies,
- multimedia tools, and Internet applications for educational purposes has led to the emergence of
- a new literacy context in which specific cognitive processes are required for effective reading
- comprehension (Anderson, 2001; Chun & Plass, 1997; Coiro & Dobler, 2007; Spiro, Klautke,
- and Johnson 2015; Trackman et al., 2019). Coiro (2003) observes that using electronic texts
- means that new supports as well as new challenges will have a great impact on the L2 learners'
- ability to comprehend what they read. According to Park & and Kim (2016), CBR "requires
- readers to adapt their text-based reading strategies, and it brings both challenges and

هذه الطبعة إهـداء من المجمـع
ولا يُسمح بنشرها ورقياً أو تداولها تجارياً

Introduction

In recent decades, increased attention on the processes used in learning English as a foreign language (EFL) has triggered interest in examining EFL testing practices from the perspective of the test takers themselves (Cheng & DeLuca, 2011; Cohen & Upton, 2007; Huang, 2016; O'Malley & Chamot, 1990; Oxford, 1990; Xie, 2015; Gu, et al., 2015). The examination of L2 learners' test-taking strategies is directly connected to the test validity framework, within which both the processes that test takers adopt and the outcomes are assigned equal importance (Bachman, 1990; Cohen, 1996; Cohen & Upton, 2007; Hirano, 2009). Test-taking strategies are the cognitive language learning processes that L2 students have selected and consciously used, at least to some degree, to answer specific language test questions (Anderson et al., 1991; Cohen, 1998'; Waiprakhon and Jaturapitakkul, 2018). A crucial part of test validation involves gathering precise information on test takers' behaviours and their specific approach when responding to test questions (Bvuma & Adika, 2017; Cohen, 2006; Hirano, 2009; Weir, 2005). Weir (2005) believes that "what a task actually tests is what is central to establishing validity" (p. 243). Research on test-taking strategies aims to determine the correlation between thought processes and assessment tools. There exists a "growing recognition of the importance of gaining a better understanding of how … test-taking strategies are used on tests as part of the process of construct validation" (Cohen & Upton, 2007, p. 209).

A major goal of language testing specialists is to gain a comprehensive understanding of the cognitive processes EFL learners use, part of which is better understanding of the processes that occur during test-taking situations (Anderson, 2001; Bachman, 2000). Moreover, with the advent of computer-based testing (CBT), which differs from the conventional paper-based testing (PBT) format in terms of the constructs to be measured, the significance of examining such strategies has become crucial (Yeom & Jun, 2020; Sawaki, 2001). Reading from a screen is a key component of CBT that has attracted "the attention of researchers, language

مستخلص

حظيت الاستراتيجيات المعرفية المُوظفةُ في الإجابة على أسئلة الاختبارات باهتمام كبير الأربعين عامًا الماضية. وهناك كمٌّ مُعتبرٌ من الأدبيات البحثية عن عديدٍ من المتغيرات ذات الارتباط باستراتيجيات الإجابة على الأسئلة، غير أن التحليلات الدقيقة لهذه الاستراتيجيات التي تلجأ إليها مجموعات معينة من الطلاب فيما يتعلق بالصيغة التي يُقدّمُ بها محتوى الاختبار لم تُسْتَقصَ بحثياً بشكل كافٍ. إن تفحصاً أعمق في هذا المجال سيعود بالفائدة على الباحثين والمعلمين على السواء، خاصة مع تنامي الاعتماد على التدريس والتقييم باستخدام الحاسوب. وتعرض هذه الورقة نتائج دراسة تستكشف تأثير تباين صيغة عرض المادة المطلوب قراءتها (قراءة عبر شاشة الحاسوب وقراءة ورقية) على استراتيجيات الإجابة على الاختبارات التي يستخدمها متعلمو اللغة الإنجليزية باعتبارها لغةً أجنبية في المملكة العربية السعودية عند الإجابة على أسئلة الاختيار من متعدد لقياس الفهم القرائي. وتبين من الدراسة أنه تم وبشكل فارقٍ إحصائياً استخدام عدد أكبر من استراتيجيات الإجابة في حال القراءة عبر الشاشة مقارنة بالقراءة الورقية. وعلاوة على ذلك، استخدم الطلاب الذين يقرؤون من الشاشة وبشكل فارقٍ إحصائياً عدداً أكبر من استراتيجيات إجراء ا الإجابة ذات النمط التصاعدي من الأسفل إلى الأعلى. وقد تفسر مثل هذه النتائج السبب الكامن وراء ما كشفت عنه بعضُ من دراسات عن أداءٍ أفضل للخاضعين للاختبارات المحوسبة إذ يبدو أن القراءة من الشاشة تشجعهم على الرجوع إلى نصوص القراءة على عكس القراءة الورقية.

الكلمات المفتاحية:

استراتيجيات الاجابة على الاختبار، القراءة الورقية، القراءة الحاسوبية، استيعاب المقروء.

هذه الطبعة إهـداء من المجمـع
ولا يُسمح بنشرها ورقياً أو تداولها تجارياً

اختبارات اللغة
Journal of Language Testing

How Different Are Test-Taking Strategies with Computer-Based and Paper-Based Reading? The Case of Saudi EFL Learners

د.عبدالعزيز بن عبدالرحمن أبانمي[١]

Abstract

The cognitive strategies used in test-taking situations have received significant attention within the past 40 years. Considerable literature is available on many of the variables related to test-taking strategies; however, finer-grained analyses of the test-taking strategies that specific student populations employ in relation to the format in which test materials are presented is still under-researched. More in-depth investigation would be helpful to researchers and educators, especially given the increased reliance on computer-mediated teaching and assessment. This paper presents the results of a study exploring the effect of variation of reading presentation format (screen and paper-based reading) on the test-taking strategies of Saudi Arabian English-as-a-foreign-language learners when answering multiple-choice reading comprehension questions. The study indicates that significantly more test-taking strategies, particularly bottom-up strategies, were mobilized with computer-based than with paper-based reading. Such findings may explain why some previous studies have found better performance among computer-based test takers since reading from a screen seems to encourage test takers to refer more to the reading texts in contrast to paper reading.

Keywords:

test-taking strategies, paper-based reading, computer-based reading, reading comprehension

(١) أستاذ مشارك في كلية اللغات والترجمة بجامعة الملك سعود

مجمع الملك سلمان
العالمي للغة العربية
King Salman Global Academy for Arabic Language

مجمع الملك سلمان العالمي للغة العربية
King Salman Global Academy for Arabic Language

اختبارات اللغة
Journal of Language Testing

هذه الطبعة
إهداء من المجمع
ولا يُسمح بنشرها ورقياً أو تداولها تجارياً

ردمد: ١٦٥٨-٩٤٩١